Title

Transcriptional network governing extraembryonic endoderm cell fate choice



Authors

Pham, Paula Duyen

Lu, Hanbin

Han, Han

et al.

Publication Date

2023-10-01

DOI

10.1016/j.ydbio.2023.07.002

Peer reviewed

Published in final edited form as:

*Dev Biol*. 2023 October ; 502: 20–37. doi:10.1016/j.ydbio.2023.07.002.

# Transcriptional network governing extraembryonic endoderm cell fate choice

**Paula Duyen Pham**[a,1], **Hanbin Lu**[b,1], **Han Han**[a], **Jeff Jiajing Zhou**[a], **Aarushi Madan**[a], **Wenqi Wang**[a], **Cornelis Murre**[b], **Ken W.Y. Cho**[a,*]

[a] Department of Developmental and Cell Biology, University of California, Irvine, CA, 92697, USA

[b] School of Biological Sciences, Department of Molecular Biology, University of California at San Diego, La Jolla, CA, 92039, USA

## Abstract

The mechanism by which transcription factor (TF) network instructs cell-type-specific transcriptional programs to drive primitive endoderm (PrE) progenitors to commit to parietal endoderm (PE) versus visceral endoderm (VE) cell fates remains poorly understood. To address the question, we analyzed the single-cell transcriptional signatures defining PrE, PE, and VE cell states during the onset of the PE-VE lineage bifurcation. By coupling with the epigenomic comparison of active enhancers unique to PE and VE cells, we identified GATA6, SOX17, and FOXA2 as central regulators for the lineage divergence. Transcriptomic analysis of cXEN cells, an *in vitro* model for PE cells, after the acute depletion of GATA6 or SOX17 demonstrated that these factors induce *Mycn*, imparting the self-renewal properties of PE cells. Concurrently, they suppress the VE gene program, including key genes like *Hnf4a* and *Ttr*, among others. We proceeded with RNA-seq analysis on cXEN cells with FOXA2 knockout, in conjunction with GATA6 or SOX17 depletion. We found FOXA2 acts as a potent suppressor of *Mycn* while simultaneously activating the VE gene program. The antagonistic gene regulatory activities of GATA6/SOX17 and FOXA2 in promoting alternative cell fates, and their physical co-bindings at the enhancers provide molecular insights to the plasticity of the PrE lineage. Finally, we show that the external cue, BMP signaling, promotes the VE cell fate by activation of VE TFs and repression of PE TFs including GATA6 and SOX17. These data reveal a putative core gene regulatory module that underpins PE and VE cell fate choice.



* Corresponding author. kwcho@uci.edu (K.W.Y. Cho).

[1] These authors contributed equally.

Author contributions

Conceptualization & Methodology: P.D.P., H.L., K.W.Y.C.; Experiment investigation: P.D.P. and H.L. with help from H.H., J.J.Z., A.M.; Software, Formal analysis, Data curation & Visualization: H.L. and P.D.P. Manuscript writing: P.D.P, H.L., K.W.Y.C.; Resources: K.W.Y.C, C.M., W.W.; Supervision: K.W.Y.C.

Declaration of interests

The authors declare no competing interests.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ydbio.2023.07.002.

**Keywords**

Parietal endoderm; Visceral endoderm; Primitive endoderm; Transcription factors; Degron tags; BMP signaling; Gata6; Sox17; FoxA2; Mycn

## 1. Introduction

The amnion, chorion, yolk sac, and allantois are critical extraembryonic membranes present in mammals, and proper extraembryonic tissues are essential for nutritive support and survival of fetus as well as assisting the patterning of the embryonic axes (Ferner and Mess, 2011; Sheng and Foley, 2012; Filimonow and Fuente, 2021). The origin of extraembryonic tissues can be traced back to the blastocyst stage of embryonic development. This is when trophectoderm (TE) cells and the inner cell mass (ICM) cells develop from the totipotent progenitors that co-express CDX2 and POU5F1 (OCT4) (Niwa et al., 2005; Dietrich and Hiiragi, 2007; Holmes et al., 2017). The status quo of the progenitor cell state is disrupted by the differential activation of Hippo signaling between outer and inner cells, determined by the position-dependent polarities formed following cell cleavages (Nishioka et al., 2009; Leung and Zernicka-Goetz, 2013; Rayon et al., 2014). In outer cells, inactivation of Hippo signaling allows nuclear translocation of transcription coactivator YAP, enabling it to interact with TEAD4. The formation of YAP/TEAD complexes in the nucleus promotes *Cdx2* expression and suppresses *Oct4* expression, thus committing the outer cells to the TE cell fate (Nishioka et al., 2009; Leung and Zernicka-Goetz, 2013; Rayon et al., 2014). By contrast, the activation of Hippo signaling in the inner cells sequesters YAP to the cytoplasm, thereby preventing YAP/TEAD complex formation. Consequently, CDX2 expression is lost, but OCT4 expression increases. This ultimately drives the inner cells to the ICM cell fate (Nishioka et al., 2009; Leung and Zernicka-Goetz, 2013; Rayon et al., 2014).

Whereas TE cells gives rise to the extraembryonic structures, the ICM cells develop to establish the embryo proper (Filimonow and Fuente, 2021). Initially, all ICM cells co-express NANOG and GATA6, lineage determinant factors for epiblast (EPI) and primitive endoderm (PrE) (Chazaud et al., 2006; Bessonnard et al., 2014; Yamanaka et al., 2010; Molotkov et al., 2017; Kang et al., 2013, 2017; Hamilton et al., 2019). Subsequently, EPI-biased cells, characterized by elevated NANOG expression and decreased GATA6 expression, as well as PrE-biased cells that display the opposite NANOG-GATA6 expression pattern, emerge. These cells are interspersed throughout the ICM in a ‘salt-and-pepper’ manner (Chazaud et al., 2006; Plusa et al., 2008; Schrode et al., 2014; Bessonnard et al., 2014; Xenopoulos et al., 2015; Saiz et al., 2016; Kang et al., 2013). EPI-biased cells secrete FGF4 ligand into the ICM environment (Chazaud et al., 2006; Plusa et al., 2008; Schrode et al., 2014; Bessonnard et al., 2014; Xenopoulos et al., 2015; Saiz et al., 2016; Kang et al., 2013). In response to FGF signaling, PrE-biased cells maintains GATA6 expression, leading to the fully-committed PrE cells (Chazaud et al., 2006; Plusa et al., 2008; Schrode et al., 2014; Bessonnard et al., 2014; Xenopoulos et al., 2015; Saiz et al., 2016; Kang et al., 2013). Conversely, the insensitivity of EPI-biased cells to the FGF4 ligand reinforces NANOG expression, thereby driving the commitment to the EPI cell fate

(Chazaud et al., 2006; Bessonnard et al., 2014; Yamanaka et al., 2010; Molotkov et al., 2017; Kang et al., 2017; Hamilton et al., 2019). By undergoing these two waves of cell fate bifurcations, the early embryo develops three distinct and spatially-organized lineages - TE, EPI, and PrE (Fig. 1A). Notably, a shared approach in these cell fate decisions involves the utilization of developmental cues (in this case Hippo and FGF signaling) to segregate the bipotential progenitor into two underlying cell types by controlling the expressions of master transcription factors (TFs).

During the late blastocyst stage, PrE cells differentiate to become parietal endoderm (PE) and visceral endoderm (VE) cells (Filimonow and Fuente, 2021), which are major subtypes of the extraembryonic endoderm (ExEn) (Fig. 1A). Previous studies reported that VE cells in post-implantation embryos can undergo transdifferentiation to PE cells (Cockroft and Gardner, 1987; Casanova and Grabel, 1988; Ninomiya et al., 2005; Yagi et al., 2016), while a PE-like extraembryonic endoderm (XEN) cell line acquires VE identity in response to BMP signaling (Artus et al., 2012; Paca et al., 2012; Sozen et al., 2018; Zhang et al., 2019). This plasticity in mature PE/VE cells suggests that differentiating PrE cells are in a bipotential state that can adopt PE or VE cell fates. While previous genetic knockout studies implied the critical roles of GATA (GATA4 and GATA6) and SOX (SOX7 and SOX17) TFs in the adoption of PE cell fate (Schrode et al., 2014; Koutsourakis et al., 1999; Kanai-Azuma et al., 2002; Shimoda et al., 2007; Cai et al., 2008; Lim et al., 2008; Qu et al., 2008; Morris et al., 2010; Niakan et al., 2010; Artus et al., 2011; McDonald et al., 2014; Wamaitha et al., 2015), it is often difficult to interpret lineage-specific functions due to their earlier essential roles as well as exhibiting overlapping expression patterns in many different tissue types and TF redundancy (Schrode et al., 2014; Koutsourakis et al., 1999; Kanai-Azuma et al., 2002; Cai et al., 2008; Artus et al., 2011; Molkentin et al., 1997; Kuo et al., 1997; Bowles et al., 2000; Wat et al., 2012). Moreover, there is the challenge in accessing scarce material of pre-to peri-implantation embryos for contemporary genomic studies. Therefore, the molecular mechanism underlying this decision-making process has remained obscure. Using a combined bioinformatics and cell-based approaches, our goal is to uncover the critical TF regulatory network and potential signaling cues that control the generation of PE and VE cells.

Here, our study focused on analyzing the transcriptionally defined PrE, PE, and VE cell states based on scRNA-seq dataset of E4.5 mouse embryos (Mohammed et al., 2017; Cheng et al., 2019; Nowotschin et al., 2019; Qiu et al., 2022). Through comparative analysis of enhancer landscapes specific to PE and VE cells, we discovered that a subset of active enhancers present in VE cells was also highly accessible in PE cells. DNA motif enrichment analysis suggested that these enhancers were regulated by interaction involving FOXA2, GATA6 and SOX17. To directly address the roles of these TFs in PE and VE specification, we developed XEN cells harboring a dTAG-inducible $FKBP12^{F36V}$ degron system (Nabet et al., 2018), which allowed for rapid degradation of tagged GATA6 and SOX17 proteins in combination with *FoxA2* knockout. Our findings suggest that PrE cells possess bipotential characters, and that the decision to differentiate into either PE or VE is influenced by an interplay between GATA6, SOX17, and FOXA2, as well as BMP signaling.

## 2. Results

### 2.1. Gene signatures specifying parietal endoderm (PE) and visceral endoderm (VE) cell fates

To investigate the cell states within the extraembryonic endoderm (ExEn) lineage at embryonic day 4.5 (E4.5), we integrated published scRNA-seq datasets of E4.5 embryos (Mohammed et al., 2017; Cheng et al., 2019; Nowotschin et al., 2019; Qiu et al., 2022), followed by cell clustering analysis. We detected TE, EPI, and PrE (ExEn) lineages as reported previously (Fig. 1A and B, left panel). Furthermore, the integrated dataset with a greater number of single cells (311 cells) enabled us to detect three subpopulations within the ExEn lineages (Fig. 1B, left panel). The analysis of the expression levels for *Pou5f1* (*Oct4*), *Lama1*, and *Ttr* (markers of PrE, PE, and VE) (Niimi et al., 2004; Kwon and Hadjantonakis, 2009; Wallingford et al., 2013; Ohinata et al., 2022) in these cell subpopulations revealed characteristic expression patterns: $Oct4^{hi}Lama1^{med}Ttr^{med}$, $Oct4^{med}Lama1^{hi}Ttr^{lo}$, and $Oct4^{lo}Lama1^{lo}Ttr^{hi}$, corresponding to the PrE, PE, and VE cell states (Figs. S1A and S1B). Single cell trajectory inference reconstructed from E4.5 and E5.5 embryos supported the marker-based assignment of these cell states (Fig. 1A and B, right panel and S1C). The candidate PrE cell subpopulation was positioned at the root of ExEn lineage bifurcations, which diverged to the PE and VE cell subpopulations (Fig. 1A and B, right panel and S1C). These cells further differentiated to form E5.5 PE and E5.5 embryonic VE (emVE) and extraembryonic VE (exVE) cells (Fig. 1A and B, right panel and S1C).

To identify transcriptional signatures that distinguish PE and VE cell states at the E4.5 stage, we performed high-dimensional weighted gene correlation network analysis (hdWGCNA) (Langfelder and Horvath, 2008; Morabito et al., 2021). The resulting PE and VE gene modules were preferentially expressed in their respective PE and VE cells at both E4.5 and E5.5 stages (Fig. 1C, Tables S1 and S2). Gene ontology (GO) analysis on the mouse anatomical ontology database (Hayamizu et al., 2013) showed that these modules were related to PE or VE tissue terms (Fig. S1D). Moreover, KEGG pathway analysis (Kanehisa and Goto, 2000) also showed enrichment for biological functions related to PE or VE cells (Filimonow and Fuente, 2021) (Fig. 1D and E). Specifically, the primary pathways of the PE module, such as “protein processing in endoplasmic reticulum”, “focal adhesion”, “ECM-receptor interaction”, and “proteoglycans in cancer”, were associated with production and processing of the basement proteins (Adamson and Ayers, 1979; Cooper et al., 1981; Lane et al., 1983; Salamat et al., 1993, 1995; Semoff et al., 1982; Ueda et al., 2020; Smyth et al., 1999). These basement proteins form a specialized basement membrane, Reichert’s membrane, that structurally supports the embryo (Adamson and Ayers, 1979; Cooper et al., 1981; Lane et al., 1983; Salamat et al., 1993, 1995; Semoff et al., 1982; Ueda et al., 2020) (Fig. 1D). Interestingly, we also identified pathways related to self-renewal and stemness such as “PI3K-Akt signaling pathway” and “signaling pathways regulating pluripotency of stem cells” (Fig. 1D). The self-renewal property of *in vivo* PE cells were consistent with the successful establishment of ex vivo PE-like, rather than VE-like, cell models (Ohinata et al., 2022; Niakan et al., 2013; Cho et al., 2012; Strickland and Mahdavi, 1978; Strickland et al., 1980). Distinguishably, GO and KEGG terms enriched for the VE module were “lysosome”,

"endocytosis", "fatty acid metabolism", "steroid biosynthesis", and "mineral absorption". These terms align with the known nutritional support functions of VE cells that facilitate biosynthesis, processing, and transport of nutrient, hormone, apolipoproteins, and other molecules from the maternal environment, thereby supporting embryonic cell development (Filimonow and Fuente, 2021; Meehan et al., 1984; Donovan et al., 2005; Kawamura et al., 2012; Perea-Gomez et al., 2019; Kawamura et al., 2020; Sun-Wada et al., 2021; Sun-Wada and Wada, 2022; Assmat et al., 2005; Yu et al., 2014) (Fig. 1E). Next, we constructed molecular interaction networks for both gene modules to annotate the enriched KEGG pathways at the gene level (Fig. 1D and E). This analysis revealed that *Myc* connects the self-renewal and stemness pathways of PE module (Fig. 1D, see details below), whereas VE module was enriched for gene families such as *V-ATPases*, *Rab*, *Slc*, and other genes related to VE's nutritional support function to ensure fetal survival (Filimonow and Fuente, 2021; Meehan et al., 1984; Donovan et al., 2005; Kawamura et al., 2012; Perea-Gomez et al., 2019; Kawamura et al., 2020; Sun-Wada et al., 2021; Sun-Wada and Wada, 2022; Assmat et al., 2005; Yu et al., 2014) (Fig. 1E, see details below). Taken together, we uncovered distinct transcriptional programs associated with PE and VE cell states that diverge in E4.5 embryos.

## 2.2. Primitive endodermal cells (PrE) express TFs linked to PE and VE differentiation

To investigate the transcriptional state of PrE progenitors, we projected the single-cell transcriptomes of ExEn cells at E4.5 stage into a two-dimensional space defined by the transcriptional signatures of PE and VE gene modules (Fig. 2A). The positioning of PrE cells between the PE and VE cell clusters suggests that PrE cells exhibit intermediate expression of both PE and VE gene modules (Fig. 2A, Tables S1 and S2). The genes exhibiting the highest correlation (top 5%) within the PE and VE modules were evenly expressed in PrE cells, yet their expressions were notably biased towards either PE or VE cells (Fig. 2B, Tables S1 and S2).

Since TFs are known to control gene expression programs, we then examined TFs that were highly correlated with PE or VE modules (Fig. S2A). We identified TF sets that were preferentially expressed (hdWGCNA analysis) (Langfelder and Horvath, 2008; Morabito et al., 2021) in the single cells of PE and VE cell states (Fig. S2A, Tables S3 and S4). The gene expression, which were averaged from the identified cell populations (see Fig. 1B, left panel), of the top correlated TFs were strongly biased toward either PE or VE cells (Fig. 2C and D; Tables S3 and S4). Interestingly, many of these PE- or VE-associated TFs were already expressed in PrE cells, albeit at lower levels (Fig. 2C and D). These findings suggest that progenitor PrE cells inherently exhibit the transcriptional programs for PE and VE cells.

## 2.3. Regulation of VE cell fate mediated by Gata6 and Sox17

To interrogate how 'cell-type-specific' TFs instruct the PE or VE gene programs, we generated a cXEN cell line (chemically induced PE-like extraembryonic cells derived from mESCs) (Niakan et al., 2013; Cho et al., 2012). The derived cXEN cells were PDGFRA positive (a cell surface marker of PrE/PE ExEn cells) (Plusa et al., 2008; Cho et al., 2012; Artus et al., 2010) (Fig. S2B). RNA-seq profiling on these cells show that they expressed key PE marker genes (*Lama1*, *Lamb1*, *Col4a1*, *Col4a2*) (Adamson and Ayers, 1979; Alpy et al., 2005; Haniel et al., 1995; Miner et al., 2004) and resembled to *in*

*vivo* (extraembryonic) PE cells (Fig. 1D) (Figs. S2C and S2D). By performing ATAC-seq and H3K27ac CUT&RUN on cXEN cells and comparing them to matching epigenetic profiles (GSE125318) derived from E6.5 VE cells (Xiang et al., 2020), we identified unique and shared enhancers in PE and VE cell types. Enhancer regions, identified based on the enrichment of H3K27ac but lacking the deposition of H3K4me3 in cXEN and E6.5 VE cells, were clustered according to the joint signals of H3K27ac and the chromatin accessibility and visualized by a tornado plot (Fig. 2E). Then, active enhancers specific for PE or VE cells were linked to nearby genes (Table S5). We found that PE- and VE-enhancer nearby genes were upregulated in the corresponding cell types at the E4.5 and E5.5 stages (Fig. S2E, Table S5). PE enhancers were enriched for GATA, SOX, KLF, TEAD, and JUN TF motifs, which coincided with the list of TFs strongly correlated with the single-cell PE module (Fig. 2C and F, Table S3). Unexpectedly, VE enhancers were highly enriched for the PE-associated GATA and SOX TF motifs (Fig. 2C and F, Table S3), as well as FOX, HNF1, HNF4, and LHX motifs, whose members (FOXA2, HNF1b, HNF4a, and LHX1) were found in the VE module (Fig. 2D and F, Table S4). The enrichment of PE-associated GATA and SOX cis elements in VE enhancers suggests that VE enhancers were co-regulated by both PE- and VE-associated TFs.

Surprisingly, VE enhancers in PE-like cXEN cells, were accessible, albeit inactive (note the lack of H3K27ac), indicative of TF occupancies (Fig. 2E). Therefore, we examined the occupancies of GATA6, SOX17, and FOXA2 (GS&F) on the VE enhancers in cXEN cells, as these are the motifs most enriched in these enhancer regions (Figs. 2F, 3A and S3A). Our CUT&RUN data revealed that GATA6 and SOX17 were colocalized with FOXA2 at a subset of VE enhancers (30.3%), which we termed GS&F-co-bound VE enhancers (Fig. 3A and S3A). These GS&F-co-bound VE enhancers possessed higher levels of H3K27ac and chromatin accessibility when comparing to the remaining VE enhancers that were unbound by GATA6, SOX17 and FOXA2 in cXEN cells (Fig. 3A and S3A–C). Nevertheless, these levels were still significantly lower than those of active PE enhancers in cXEN cells (Fig. 2E). Thus, the GS&F-co-bound VE enhancers in cXEN cells were open but unable to be activated, signifying a “latent” state. To identify gene targets of GS&F-co-bound VE enhancers, we filtered the enhancer-proximal genes (Fig. S3D, top panel, and Table S6) that were not differentially expressed when comparing the bulk RNA-seq data between E6.5 VE and E6.5 epiblast cells (GSE76505) (Zhang et al., 2018). The remaining genes, denoted as GS&F-co-bound VE enhancer targets, were preferentially expressed in VE cells (Fig. 3B, Table S7). Noticeably, these genes were expressed at significantly higher levels in extraembryonic VE (exVE), indicating that the exVE may represent the default cell fate of VE cells, as previously proposed (Fig. 3B, Table S7) (Nowotschin et al., 2019). Taken together, we propose that the VE transcriptional program is jointly regulated by *Gata6*, *Sox17*, and *FoxA2*.

### 2.4. Combinatorial activities involving GATA6, SOX17 and FOXA2 instruct VE enhancer activity

To better understand the enhancer activity of the GS&F-co-bound VE enhancers in (PE-like) cXEN cells, we compared the accessibility profiles between the latent GS&F-co-bound VE enhancers and active PE enhancers using deeply-sequenced (~150 million reads) ATAC-seq

data (Fig. S3E). A Vplot displaying ATAC-seq fragment distribution demonstrated that active PE enhancers were highly open, as ATAC-seq fragments were spread across the entire enhancer platform (Fig. S3E, left panel). In contrast, TF occupancy was constrained at the central regions of GS&F-co-bound VE enhancers, as revealed by a well-protected region by ATAC-seq (Fig. S3E, right panel). How do GS&F-co-bound VE enhancers assume a latent state in PE cells but become active in VE cells? To address this question, we examined if there were alterations in the TF occupancies associating with GS&F-co-bound VE enhancers between VE- and PE-associated cell types using their ATAC-seq profiles. We used TOBIAS (Transcription factor Occupancy prediction By Investigation of ATAC-seq Signal) software (Bentsen et al., 2020) to compute the TF footprinting scores, which quantify the local depletion (protected by TF binding) of ATAC signal at the TF motifs relative to the flanking regions. We then performed a differential analysis on the TF footprinting scores derived from PE and VE cells (Fig. 3C, Table S8). When comparing the footprinting scores at FOXA2, HNF1B, and HNF4A motifs between E6.5 VE and (PE-like) cXEN cells, the differences were nonsignificant (Fig. 3D, Table S8). These data suggest that the VE-associated TFs (FOXA2, HNF1B, HNF4A) maintained the accessibilities of GS&F-co-bound VE enhancers regardless of latent or active states.

On the contrary, footprinting scores for GATA6, SOX17, TEAD, and JUN/FOS were significantly higher in the latent enhancer state in (PE-like) cXEN cells compared to when these enhancers are activated in E6.5 VE cells (Fig. 3D, Table S8). Indeed, this is supported by lower expression levels of *Gata6* and *Sox17* in VE cells relative to (PE-like) cXEN (Fig. 3E). Additionally, in E6.5 VE cells, there is a loss of GATA6 and *SOX17* footprinting, and a gain of (VE-associated) LHX footprinting (Fig. 3E, Table S8). These data collectively suggest a gene regulatory circuit that regulates the activation of the VE gene program. The circuit involves the repression of the VE gene program by the PE-associated TFs GATA6 and SOX17. Conversely, positive inputs from TFs FOXA2, HNF1B, HNF4A and LHX1 are required to activate the VE gene program.

### 2.5. GATA6 and SOX17 promote PE development via repression of the VE gene program and Mycn activation

GATA6 and SOX17 appear to have a dual function - positively promoting the PE cell state, while repressing the VE cell state. To examine the roles of GATA6 and SOX17 in PE cells, cXEN cells were subjected to targeted protein degradation using a dTAG-inducible FKBP12$^{F36V}$ degron system (Nabet et al., 2018) (Fig. 4A and S4A). We generated homozygous knock-in *Gata6-FKBP12*$^{F36V}$ and *Sox17*-FKBP12$^{F36V}$ cXEN cell lines (Fig. 4A and S4A). The effectiveness of the degron system was examined by treating the cells with dTAG$^{V}$-169 followed by both flow cytometry and western blotting. We observed rapid degradation of the tagged GATA6 and SOX17 proteins upon dTAG$^{V}$-1 treatment (Figs. S4B and S4C).

Next, we conducted time-course (0–96 h) RNA-seq experiments to monitor the transcriptomic alterations after GATA6 or SOX17 depletion (Fig. 4B). Heatmap clustering analysis of RNA-seq samples showed major transcriptomic changes occurring after 48 h of GATA6 or SOX17 depletion (Fig. S4D). Interestingly, the clustering analysis also

revealed that the transcriptomic changes in cXEN cells depleted for GATA6 and SOX17 shared similarity, suggesting that these two TFs regulate common pathways (Fig. S4D). Then, we performed gene set enrichment analysis (GSEA) (Subramanian et al., 2005) to assess whether a given gene set displays a statistically significant skewness towards the most affected genes by the treatment. Specifically, VE-associated gene sets were evaluated against the ranked transcriptomes where genes are ordered by the levels of gene expression alterations occurring at 72 h after the $dTAG^V$-1 treatment. This analysis uncovered that both GATA6 and SOX17 degradation upregulated the VE gene modules identified from scRNA-seq data and the GS&F-co-bound VE enhancer targets (Fig. 4D, Tables S7 and S9).

To corroborate the affected gene programs after GATA6 and SOX17 depletion, we performed Weighted Gene Co-expression Network Analysis (WGCNA) (Langfelder and Horvath, 2008), a method designed to identify co-expressed genes that are likely under the regulation of the same pathways, hence performing similar functions. The resulting heatmap was grouped into four gene modules, which we denoted as: GATA6/-SOX17 (GS) Co-upregulated genes, GATA6-Only Upregulated genes, SOX17-Only Upregulated genes, and GS Co-downregulated genes (Fig. 4B, Table S10). Notably, the marker gene for VE cells, *Ttr*, appeared in the GS Co-upregulated module (Fig. 4B). GO analysis demonstrated that GS Co-upregulated module was enriched for terms related to VE biological functions (Filimonow and Fuente, 2021; Meehan et al., 1984; Donovan et al., 2005; Kawamura et al., 2012; Perea-Gomez et al., 2019; Kawamura et al., 2020; Sun-Wada et al., 2021; Sun-Wada and Wada, 2022; Assmat et al., 2005; Yu et al., 2014) including nutritional supports (e.g., endocytosis and lipid metabolism), vasculogenesis (e.g., vasculature development and blood vessel migration), and BMP signaling (Fig. 4C). Also, GATA6-Only and SOX17-Only Upregulated modules were enriched for VE-associated terms (Fig. 4C). In sum, GATA6 and SOX17 function to suppress the VE transcriptional program.

Finally, GO analysis of GS Co-downregulated module showed enrichment for GO-terms involved in cell cycle regulation and biosynthesis (Fig. 4C). These terms were related to the functions of *Myc*-mediated pathways, which connect to the PE cell activities revealed by single-cell KEGG pathway analysis (Fig. 1D). To further determine that *Myc* activity was regulated by GATA6 and SOX17 in PE cells, we examined whether MYC target gene sets were enriched against the ranked transcriptomes affected by GATA6 or SOX17 depletion, as described earlier using GSEA (Subramanian et al., 2005). This analysis showed that the ‘Hallmarks of MYC targets’ (from the Molecular Signature Database, MsigDB) (Subramanian et al., 2005) were negatively regulated when either GATA6 or SOX17 was depleted (Fig. 4D). Interestingly, *Myc* expression was downregulated in the SOX17-depleted condition, but the expression of *Mycn* declined in both GATA6-and SOX17-depleted conditions (Fig. S4E). It is worth noting that *Mycn* was previously identified as a PE-associated TF based on scRNA-seq transcriptional signature analysis (Fig. 2C, Tables S1 and S3). Collectively, these data indicate that GATA6 and SOX17 reinforce PE-cell identity by simultaneously promoting self-renewal properties via activation of *Mycn* and *Myc* expression, while repressing a VE gene program.

### 2.6. FOXA2 maintains the activation potential of VE gene program and represses PE-associated Mycn expression

FOXA2 is a well-known pioneer TF regulating definitive and visceral endoderm development (Dufort et al., 1998; Cirillo et al., 2002; Burtscher and Lickert, 2009; Weinstein et al., 1994; Ang and Rossant, 1994). Our current data showed upregulation of *FoxA2* expression when PrE cells differentiate into VE cells (Fig. 2D), as well as the physical binding of FOXA2 to the accessible VE enhancers enriched for FOX motif in cXEN cells (Figs. 2F, 3A and S3C). These lines of evidence suggest that *FoxA2* plays a fundamental role in the regulation of VE cell fate during PE-VE lineage divergence. Therefore, to determine how the VE gene program was affected by FOXA2 in cXEN cells, we engineered a deletion mutation in the DNA binding domain of *FoxA2* gene, using a CRISPR/-Cas9 targeting method This mutation was engineered into the *Gata6*-FKBP12$^{F36V}$ or *Sox17*-FKBP12$^{F36V}$ cXEN cells (Fig. 5A). These modified cXEN cells expressed a mutant form of FOXA2 as confirmed by western blot and allowed us to simultaneously examine loss-of-function phenotypes of both FOXA2 and GATA6/SOX17 upon treatment with dTAG$^{V}$-1 (Figs. S5A and S5B). We then conducted RNA-seq analysis on the *FoxA2* mutant cell lines.

We first examined the impact of FOXA2 on cXEN cells. RNA-seq data collected from both *FoxA2*-knockout (KO) and WT cXEN cells were subjected to Differentially Expressed Genes (DEGs) analysis using DESeq2 (Table S11) (Love et al., 2014). Several VE-associated genes such as *Ttr* and *Hnf4a* were downregulated, while *Mycn* was upregulated (Fig. 5B). This data suggests that the gene regulatory direction of FOXA2 is inverse to that of GATA6 or SOX17. The downregulation of VE-associated genes resulting from the FOXA2 deletion mutant prompted us to explore how GATA6, SOX17, and FOXA2 co-bound (GS&F-co--bound) VE enhancer targets (identified previously in Table S7 and Table S9) were impacted. We used GSEA (Subramanian et al., 2005) to analyze the enrichment for these target genes against the transcriptome, ranked by the levels of expression changes between *FoxA2*-KO to WT cXEN cells. In fact, GS&F-co-bound VE enhancer targets exhibited a bias towards the downregulation end within the *FoxA2*-KO transcriptome. This suggests that FOXA2 has an activation effect and is required to sustain the basal expression levels of the GATA6/SOX17-repressed VE gene program in cXEN cells (Fig. 5C).

To examine the interplay between FOXA2 and GATA6 or SOX17, we cultured *Gata6*-FKBP12$^{F36V}$/*FoxA2*-KO and *Sox17*-FKBP12$^{F36V}$/*FoxA2*-KO cXEN cells with and without dTAG$^{V}$-1, and then analyzed the RNA-seq. All of the up-regulated and down-regulated gene sets that were altered by GATA6 or SOX17 removal (see Fig. 4B) were affected by *FoxA2* deletion (Fig. 5D and S5D). Specifically, *FoxA2* deletion prevented the activation of GATA6/SOX17 Co-upregulated, SOX17-Only Upregulated, and GATA6-Only Upregulated genes (Figs. 4B, 5D and S5D). Therefore, transcriptomics of cXEN cells lacking both GATA6 or SOX17 and functional FOXA2 were very similar to that of *FoxA2*-KO cells (Fig. S5C). These data underscore the notion that FOXA2 maintains the activation potential of the VE gene program in (PE-like) cXEN cells upon eviction of GATA6 or SOX17 from VE enhancers.

On the other hand, *Mycn* expression was significantly increased in *FoxA2*-KO cXEN cells (Fig. 5E). In the absence of GATA6 or SOX17 in *FoxA2*-KO, cXEN cells were

no longer able to downregulate *Mycn* expression or repress the (*Mycn*-associated) GATA6/SOX17 co-downregulated module (Figs. 4B, 5D and 5E). Concomitantly, MYC target genes (Subramanian et al., 2005) were induced as revealed by GSEA analysis (Fig. 5F). Thus, FOXA2 is a potent repressor for *Mycn*, whose activity is further tuned by GATA6 and SOX17, potentially preventing abnormal PE cell growth. Taken together, our transcriptomic analyses on the combined loss-of-function of FOXA2/GATA6 or FOXA2/SOX17 in cXEN cells indicate that FOXA2 counteracts GATA6 and SOX17 activities in regulating both the VE gene program and MYC/-MYCN pathways.

### 2.7. VE specification requires BMP signaling integrating with the Gata6, Sox17, Mycn, and FoxA2 network

GO term analysis of GATA6 and SOX17 targets (Fig. 4C) identified BMP signaling pathways as critical components affected after GATA6/SOX17 depletion in cXEN cells (Figs. 4C and 6A). GSEA analysis showed that both canonical (mediated by SMAD proteins) and non-canonical (transduced by ERK cascade) BMP signaling pathways (Derynck and Budi, 2019) were reinforced upon the loss of GATA6 or SOX17 (Fig. S6A), exemplified by the upregulation of individual BMP signaling components such as BMP ligands (*Bmp1*, *Bmp6*), a BMP receptor (*Acvr1*), *Smad1*, and multiple MAPK kinases (Fig. 6A). This suggests that GATA6 and SOX17 attenuate BMP signaling in cXEN cells.

To understand the impact of BMP signaling on PE-VE cell fate programming, cXEN cells were treated with BMP4 and the gene expression changes were measured using RNA-seq. BMP4 treatment downregulated the expression of PE-associated TFs, such as *Gata6*, *Sox17*, and *Mycn*, while it upregulated the expression of VE-associated TFs (Fig. 6B). We conducted GSEA analyses on three sets of genes – MYC hallmark gene sets (Subramanian et al., 2005) (Fig. 6C), our scRNA-based VE gene module, and GS&F-co-bound VE enhancer gene targets (Fig. S6B, Tables S7 and S9) in BMP4-treated cXEN cells. The enrichment analyses revealed that BMP4 treatment led to suppression of MYC/MYCN pathways and activation of VE-associated genes and VE-associated biological processes (Artus et al., 2012; Paca et al., 2012). These observations recapitulated the transcriptome phenotype of GATA6/SOX17-depleted cXEN cells (Fig. 4B and S6C, Tables S10 and S12). In summary, we propose that PE and VE were antagonistic cell fates whose divergence is determined by the integration of BMP signaling and a transcription factor network wired by *Gata6*, *Sox17*, *Mycn*, and *FoxA2* (Fig. 6D).

## 3. Discussion

Here we revealed the lineage segregation mechanism orchestrating the differentiation of primitive endoderm (PrE) into parietal endoderm (PE) and visceral endoderm (VE), which begins at around E4.5-E5.0. Through detailed mining of the scRNA-seq datasets of E4.5 embryos, we have defined transcriptional activities closely associated with PrE, PE, and VE cell states at the onset of PE/VE cell lineage segregation. This systematic analysis allowed us to identify known TFs involved in PrE differentiation and implicate other TFs that have not yet been linked with extraembryonic endoderm (ExEn) development. We found that PrE is a bipotential progenitor cell, and that the PrE to become PE or VE is dictated by

an interplay between GATA6, SOX17, and FOXA2 TFs, which is also modulated by BMP signaling (Fig. 7A and B).

### 3.1. TF network underlying PE and VE specification

The genetic roles of GATA6 and SOX17 in PrE development have been studied extensively. *Gata6* knockout (KO) mice are embryonically lethal and *Gata6* is required for PrE specification *in vivo*, where in its absence, ICM cells adopt an epiblast cell fate (Schrode et al., 2014; Koutsourakis et al., 1999; Cai et al., 2008). These mouse phenotypes were further corroborated as GATA6 was found to be a potent factor in reprogramming of mouse embryonic stem cells (mESCs) to ExEn cells (XEN cells) (Wamaitha et al., 2015). However, due to these early GATA6 phenotypes in PrE, the later role of GATA6 in PE-VE lineage divergence has not been well understood, though, GATA6 protein expression is lost in future VE cells and maintained in PE cells (Cai et al., 2008; Wallingford et al., 2013). SOX17 is detected at the early blastocyst stage and later expressed at higher levels in PE-destined cells as compared to VE cells (Artus et al., 2011). Despite the midgestation mortality, the ExEn specification in *Sox17* KO mice was unaffected (Kanai-Azuma et al., 2002). This suggested that SOX17 might be dispensable, potentially due to compensatory effects of other SOX factors, such as SOX7 (Kanai-Azuma et al., 2002). However, when *Sox17* KO embryos were induced into a diapause state, SOX17 was required for the maintenance of PrE lineage (Artus et al., 2011). Furthermore, overexpression of SOX17 in eight-cell stage embryos influences Epi-PrE cell fate allocation (Morris et al., 2010). Also, by merely overexpressing SOX17, mESCs can be reprogrammed to extraembryonic endodermal (ExEn) cell lines (McDonald et al., 2014). While these analyses establish the essential functions of GATA6 and SOX17 in the PrE (ExEn progenitor cells), regulatory mechanisms underlying the subsequent ExEn tissue development have remained elusive.

Here, we have built a core TF network that controls PE and VE cell fates (Fig. 7), beginning with a bioinformatics approach. Specifically, we compared the enhancer landscapes between PE-like (XEN cells) and E6.5 VE cells to identify TF binding motifs associated with PE-specific or VE-specific enhancer repertoires. Both PE and VE enhancers were enriched for GATA and SOX motifs, suggesting key roles for these TFs for PE and VE cell differentiation. A notable finding is that VE enhancers were strongly co-enriched for a FOXA2 (VE TF) binding motif in conjunction with GATA6 and SOX17 (PE TFs) binding motifs, suggesting a functional interaction in VE specification. To directly address this question, we have used a dTAG-inducible $FKBP12^{F36V}$ degron system to deplete GATA6 and SOX17 in PE-like cXEN cells and interrogated the transcriptomic alterations in the presence and absence of functional FOXA2. Specifically, we identified a TF regulatory circuit where GATA6 and SOX17 act to suppress the VE gene program in PE cells, while the activation potential is maintained by FOXA2 (Fig. 7A). We propose a TF-antagonizing model during PrE lineage divergence, where GATA6 and SOX17 promotes the PE gene program cell in PE cells, while suppressing the VE gene program in PE cells (Fig. 7A). On the contrary, when GATA6 and SOX17 effects are weakened, PrE cells differentiate toward the VE cell lineage that are dependent on FOXA2 (Fig. 7B). The mechanism by which GATA6 and SOX17 impart repression on VE enhancers, while activating PE enhancers is not well understood. One potential mechanism is that the motif combination, compared to

that without FOXA2 motif, influences the physical occupancy of GATA6 and SOX17 with FOXA2 on enhancers, thereby dictating distinct enhancer activities (known as enhancer grammatical effect) (Jindal and Farley, 2021). Alternatively, enhancer activities may be influenced by alterations in epigenetic states of PE versus VE enhancers by differential TF recruitment of epigenetic cofactors and/or remodelers.

In addition to GATA6, SOX17, and FOXA2 TF motifs, our analysis suggests the involvement of other TFs (Fig. 7). Identification of HNF1B and HNF4A motifs agrees with the previous findings for HNF1B and HNF4A requirements in VE development (Coffinier et al., 1999; Barbacci et al., 1999; Chen et al., 1994; Duncan et al., 1997). Similar to FOXA2 motif, the TF footprint of HNF1B and HNF4A motifs remained unchanged in the VE enhancers when comparing E6.5 VE to cXEN cells. It is unclear how, and whether, HNF1B and HNF4A cooperate with FOXA2 to potentiate the VE gene program. By contrast, the footprint of the LHX motif increased in E6.5 VE cells relative to cXEN. This indicates that LHX1 may be crucial to the development of advanced VE cell types, such as the anterior VE associated with embryonic VE tissues during gastrulation, as previously reported (Costello et al., 2015). Additionally, we discovered several PE TF motifs such as JUN/FOS and TEAD which have yet to be characterized in both PE-VE lineage divergence and PE development. In the future, it will be important to unveil the enhancer-regulatory roles of these TFs in context of the DNA sequence features of the VE and PE enhancers as they may manifest the evolutionary strategy that led to the emergence of ExEn tissue when the first amniotes evolved from amphibian ancestors approximately 340 million years ago (Ferner and Mess, 2011).

Our gene interference analysis revealed that *Mycn*, which is preferentially expressed in PE cells (Fig. 2C, Tables S1 and S3), is activated by GATA6 and SOX17. *Mycn* and to a lesser degree *Myc*, were concordantly upregulated in PE cells. This suggests that PE cells undergo an extensive self-renewable process (Figs. 1D and 2C, Tables S1 and S3). These findings explain why previously isolated ExEn stem cell lines (e.g. cXEN) predominantly resemble rapidly growing PrE or PE cell types (Paca et al., 2012; Ohinata et al., 2022; Kunath et al., 2005; Fowler et al., 1990). We speculate that the “stemness” properties of PE cells, likely conferred by *Myc* and *Mycn* activities, allow PE cells to maintain their cell identity, and hence they do not differentiate into more advanced cell types in the highly inductive embryo environment. Conversely, the downregulation of *Myc* factors might be crucial for VE cells to further develop into emVE/exVE cells during the E4.5 to E5.5 embryo stages, and into more specialized VE cell types during post-implantation (Fig. 7A).

FOXA2 is a dominant repressor of *Mycn*, and facilitate the adoption of VE cell fate (Fig. 7B). However, FOXA2 is also modestly expressed in the PE cells despite being a VE-associated TF. We speculate that PE cells leverage FOXA2 to potently restrict the expression of *Mycn*, which could otherwise be over-stimulated by PE TFs such as GATA6 and SOX17 (Fig. 7A). This mechanism safeguards PE cells from uncontrolled growth, mimicking FOXA2’s role as a tumor suppressor (Sahoo et al., 2022). In line with this notion, *FoxA2* KO embryos exhibit abnormal overgrowth of PrE/PE-like cells (Dufort et al., 1998), which may allude to dysregulated *Mycn* expression.

### 3.2. BMP signaling input is required for PE and VE segregation

We observed an upregulation of genes encoding BMP signaling components in cXEN cells deficient in GATA6 or SOX17 (Fig. 6A and S6A). Motivated by this observation, we investigated the genome-wide transcriptional changes in cXEN cells after BMP4 treatment. BMP4 downregulated the expression of PE-associated genes including *Gata6*, Sox17, and *Mycn* (Fig. 6B), while upregulating genes associated with VE, such as *Hnf4a* and *Hex* (Fig. 6B). This finding is consistent with the previous observation that BMP4 treatment directs XEN cells towards a VE phenotype (Artus et al., 2012; Paca et al., 2012). We propose that BMP signaling acts as a potent inducer for VE cell fate by upregulating VE TFs and by downregulating expression of PE TFs such as *Gata6* or *Sox17* (Fig. 6B). This observation also fits with the known spatial expression pattern of *Bmp4* in E4.5 blastocysts, where a high *Bmp4* expression level is detected in the epiblast cells adjacent to the future VE cells (Coucouvanis and Martin, 1999; Mochel et al., 2015; Graham et al., 2014). In contrast, the future PE cells are localized further away from the BMP source of the epiblast cells and migrate along mural trophoblast cells (mTE). Collectively, our data suggest that BMP signaling is one of the major developmental cues involved in VE versus PE cell fate decisions.

We propose the following mechanism underlying PE-VE cell specification. Initially, PrE cells moderately express *Gata6* and *Sox17*, where PrE cells are bipotential, and co-express other genes involved in PE and VE gene regulatory programs (Fig. 2A–D). The PrE cells receive BMP signals from the adjacent Epi cells, which collaborates with FOXA2 and other VE TFs to promote the VE gene regulatory program. The exact mechanism of how BMP signaling impacts the VE cell decision is currently unclear. On the other hand, the PrE cells that detach from the EPI cells, acquire PE cell fate via upregulation of *Gata6* and *Sox17*, possibly due to being further away from BMP signals, but also closer to paracrine parathyroid hormone-related protein (PHTrP) signals from adjacent mTE cells (Verheijen et al., 1999; van de Stolpe et al., 1993). We envision that recently developed primitive endoderm stem cells (PrESC) lines (Ohinata et al., 2022), which model the PrE progenitor cells, will be valuable for testing the current proposed model.

We propose that it will be crucial to distinguish the respective contributions of canonical and noncanonical BMP signaling pathways (Derynck and Budi, 2019) to the VE gene program. A potential approach is the use of a BMP response element-βgal (BRE-gal) reporter system that specifically responds to canonical (SMAD mediated) BMP signaling during pre-implantation development. This system can pinpoint the time and location of the signaling cascade in the lineage segregation (Mochel et al., 2015; Javier et al., 2012). Additionally, the precise role of BMP signaling in PE-VE cell lineage specification *in vivo* will need to be determined. Effects of BMP signals in PE/VE lineage divergence have yet to be finely delineated due to the requirement for BMP signals in both trophoblast and endoderm derived tissues (Coucouvanis and Martin, 1999; Sirard et al., 1998; Yang et al., 1998; Beppu et al., 2000; Chu et al., 2004; Sozen et al., 2021). Thus, a particularly important question is to address how modulation of BMP signaling and the interplay among GATA6, SOX17, and FOXA2 shifts the balance of PE and VE cells in developing embryos. Recent advances in the growth of integrated embryo models using mixtures of self-assembled ESC,

TSC, and XEN stem cells (Sozen et al., 2018; Zhang et al., 2019; Amadei et al., 2021, 2022; Tarazi et al., 2022), could provide a powerful approach to study the interplay among the TFs and the effect of BMP signaling in PE/VE lineage divergence. Lastly, besides demonstrating that cXEN cells are genetically editable cells, we also proved the principle that the FKBP12$^{F36V}$-based proteolysis targeting chimera (PROTAC) system is compatible with XEN cells. Treatment with dTAG$^{V}$-1 effectively recruits E3 ubiquitin ligase to the FKBP12$^{F36V}$ tagged protein followed by rapid degradation (Nabet et al., 2018; Nabet et al., 2020). The application of PROTAC in XEN or associated endodermal stem cells (Ohinata et al., 2022; Amadei et al., 2021, 2022; Tarazi et al., 2022) will expand the targetable genes for further functional studies.

In closing, we would like to emphasize that our proposed model is based on a limited number of TFs and enhancers (Fig. 7A and B). Therefore, further studies are needed to investigate the involvement of other potential regulatory factors and enhancers to gain a more comprehensive understanding of the molecular mechanism underlying extraembryonic endoderm development. However, the framework of the current gene regulatory model and insights provided by the current study can serve as a foundation for future research in understanding the transcriptional regulation of extraembryonic endoderm differentiation.

## Experimental procedures

### Lead Contact and materials availability

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Ken W.Y. Cho (kwcho@uci.edu).

## Method details

### cXEN cell line derivation

cXEN cell lines were derived from mESCs cells (E14TG2a) using a previously established chemical induction protocol (Niakan et al., 2013) in a cXEN derivation cell medium consisting of 0.010 M Retinoic Acid (Sigma, 50–185-8562), 10 ng/mL Activin A 24 ng/mL (R&D Systems, 338AC010), Fgf2 (R&D Systems, 23–3FB0–10), and 1 ug/mL Heparin (Sigma, H3393). mESCs were maintained in KnockOut DMEM (Thermo Fisher,10829018) supplied with 15% FBS (R&D Systems, S10250H), 1 mM MEM Non-Essential Amino Acids (Thermo Fisher, 11140–050), 2 mM Glutamax (Thermo Fisher, 35050061), 100 U/mL Pen/Strep (Thermo Fisher, 15140122), 0.1 mM 2(β)-ME (Sigma, M-3148), and 1000 U/mL LIF (Cell Guidance Systems, GFM200), before changing to cXEN derivation medium. The established cXEN cell lines were maintained, and passaged in Advanced RPMI (Thermo Fisher,12633012) supplemented with 15% FBS (R&D Systems, S10250H), 2 mM Glutamax (Thermo Fisher, 35050061), 100 U/mL Pen/Strep (Thermo Fisher, 15140122), 0.1 mM 2(β)-ME (Sigma, M-3148).

### CRISPR/Cas9 cell line generation

We cloned the FKBP12$^{F36V}$-mNeonGreen-(3X)HA-tag-P2A-Blasticidin cassette flanked by the homology arms of the target gene into the pUC19 backbone. The homology

arms, amplified from mouse genomic DNA or synthesized using gBlock (IDT), were designed to produce in-frame insertion at the C-terminus of the targeted proteins. SgRNAs were designed to cut around the stop codons of the endogenous loci of the *Gata6* and *Sox17* genes. The repair template and sgRNAs were validated by Sanger sequencing before use. For generation of FKBP12$^{F36V}$ cell lines, cXEN cells were seeded the day before transfection on a 10 cm$^2$ tissue culture plate. Twelve μg of plasmid DNA mixed with equimolar amounts of sgRNAs and homology repair templates were transfected using 40 μl of Lipofectamine 3000 with 24 μl of P3000 reagent (Thermo Fisher, L3000008). Two days after transfection, the transfected cells were selected in 10 μg/mL of Blasticidin for a week (Cayman Chemical, NC1445974). Single cells from the mNeonGreen positive population were then sorted into 96-well round bottom plates pre-coated with irradiated CF-1 MEFs (Thermo Fisher, A34180) using a BDFACSAria Fusion Sorter. Cell clones harboring homozygous inserts were screened by PCR genotyping using primers that detect successful insertions. The clones were verified by flow cytometry and western blots. The genotyping primers used for detecting successful knockin in the *Gata6* loci: FWD: GGTCAAGACGGCCTCTACATAGGTG, REV: CTGGTGAAATAGGTTTCCATATGATATCCC. The genotyping primers used for detecting successful knock-in in the *Sox17* loci: FWD: AGCTCAGCGGTCTACTATTGCAACTACC, REV: AGCACTCAGCA-CAGCATTTGCAGGG. The sgRNAs used for targeting *Gata6* loci are: sense: CACCGCCAAGAATCCTGTCGCACGG, antisense: AAACCCGTGCGA-CAGGATTCTTGGC. The sgRNAs used for targeting *Sox17* loci are: sense: CACCGCGGTTGCCGACCCGACCTGA, antisense: AAACTCAGGTCGGGT CGGCAACCGC.

To generate *FoxA2*-KO knockout lines on *Gata6*-or *Sox17*- FKBP12$^{F36V}$ knockin lines, two sgRNAs were designed to delete the sequence encoding the DNA binding domain. cXEN knockin cell lines were seeded one day before transfection on a 6-well tissue culture plate. A cocktail of 2.5 ug of plasmid DNA and equimolar amounts of sgRNA was transfected with 5 μl of lipofectamine 2000 (Thermo Fisher, 11668019) for 24h. The BFP positive cells were then sorted using a BDFACSAria into 96-well round bottom plates pre-coated with irradiated CF-1 MEFs (Thermo Fisher, A34180). The homozygous clones were confirmed by genotyping. The genotyping primers used for detecting *FoxA2* forkhead deletion: FWD: “AACATGTCATCC-TATGTGGGCGC” and REV: “GGTCTTCTTGCCTCCGCTACTG”. The sgRNAs used for deletion of *FoxA2*’s forkhead domain: set_1: sense: “CACCGCACA-CTTGAAGCGCTTCTGG”, antisense: “AAACCCAGAAGCGCTTCAAG TGTGC”, set_2: sense: “CACCGTGGCGTGTGTGTAGCTGCGT”, antisense: “AAACACGCAGCTACACACACGCCAC”.

## BMP4 treatment

cXEN cells were plated on a gelatin coated plate with N2B27 plus 1% FBS overnight. N2B27 media is a 1:1 mix of DMEM/F12 (Thermo Fisher, 11320033) and Neurobasal media (Thermo Fisher, 21103049) supplemented with B-27 Plus (Thermo Fisher, A3582801) and N-2 (Thermo Fisher, 17502048). The media was then changed to N2B27 without serum, supplemented with 25 ng/ml BMP4 (R&D Systems, 314-BP-010/CF). Cells are seeded and harvested at Day 1 (24h) and Day 3 (72h) for RNA-seq.

### RNA-seq

Total RNA was extracted using Trizol (Thermo Fisher, 15596026). Strand-specific RNA-seq libraries were generated with 1000 ng of total RNA using NEBNext Ultra II Directional RNA Library Prep Kit for Illumina (E7760S). Libraries were subjected to paired-end sequencing by the NovaSeq 6000.

### ATAC-seq

Fifty thousand cXEN cells were dissociated using Accutase (Thermo Fisher, A1110501). ATAC-seq was performed according to manufacturer instructions using the ATAC-seq kit (Active Motif, 53150), which included components for tagmentation, tagmented DNA clean-up, and library PCR amplification. Libraries were subjected to paired-end sequencing by the NovaSeq 6000.

### CUT&RUN

A half million cXEN cells or its derivative cells were dissociated using Accutase (Thermo Fisher, A1110501). CUT&RUN was performed using pAG-MNase (Epicypher, 15–1016) using a cell preparation method from Fujiwara et al. (Fujiwara et al., 2021), following the standard protocol published by the Henikoff lab (Skene and Henikoff, 2017; Meers et al., 2019). A half of the DNA eluted from the CUT&RUN reaction was used for library prep using the NEBNext Ultra II system based on Liu et al. (Liu et al., 2018; Liu, 2019) without size selection. Libraries were subjected to paired-end sequencing on the NovaSeq 6000. Antibodies used in these experiments were as follows: anti-GATA6 (5851S, Cell Signaling Technology), anti-SOX17 (AF1924, R&D Systems), anti-FOXA2 (07–633, EMD Millipore Sigma), anti-H3K27ac (MABE647, EMD Millipore Sigma), anti-H3K4me3 (39159, Active Motif), and anti-H3K4me1 (ab8895, Abcam).

### Flow cytometry

Cells were dissociated and resuspended in media for analyzing the expression of NeonG fluorescence protein in $FKBP12^{F36V}$ knock-in lines using the ACEA Novocyte Quanteon (Agilent). For cell surface staining, cells were dissociated, washed in 1X PBS, stained with anti-CD140A (PDGFRA) (17–1401-81, Thermo Fisher) for 15 min. Stained cells were washed twice in PBS, and resuspended in FACS buffer (PBS+1% FBS) before FACS analysis.

### Western blotting

Antibodies used for western blot analysis were as follows: anti-GATA6 (5851S, Cell Signaling Technology), anti-SOX17 (AF1924, R&D Systems), anti-FOXA2 (07–633, EMD Millipore Sigma), anti-HA (901503, Biolegend), or anti-TUBULIN (T6199, Sigma). The cells were harvested in a 1X SDS loading buffer, boiled at 95 °C before being subjected to western blotting.

## Data quantification and statistical analyses

### CUT&RUN and ATAC-seq data analysis

Paired-end fragments were trimmed using trimmomatic (Bolger et al., 2014). Paired-end reads were aligned using Bowtie2 (Langmead and Salzberg, 2012) with flags: –local –very-sensitive –dovetail –phred33 –I 10 -X 700. The fragments were filtered for a MAPQ score of 30 prior to downstream analyses. MACS2 (Zhang et al., 2008) was used for narrowPeak calling for CUT&RUN and ATAC-seq. To identify TF binding regions in H3K27ac peaks, HOMER (Heinz et al., 2010) was used for calling H3K27ac peaks centered on nucleosome free regions (NFR).

Motif enrichment analysis was performed using HOMER (Heinz et al., 2010) on cXEN and E6.5 VE H3K27ac peaks. H3K27ac-enriched enhancers were defined by filtering out H3K27ac peaks with high H3K4me3 signals through k-means clustering by Deeptools (Ramírez et al., 2014). Tornado plots and meta profiles of CUT&RUN and ATAC-seq datasets were generated by Deeptools' plotHeatmap and plotProfile subcommands (Ramírez et al., 2014). Enhancer sets for nearby genes were annotated by HOMER (Heinz et al., 2010). ATAC-seq footprinting analyses were performed using the TOBIAS pipeline with motifs supplied from the JASPAR CORE 2022 vertebrates (Castro-Mondragon et al., 2021) and Hocomoco V11 (Mouse) (Kulakovskiy et al., 2017). To plot the Vplots, VplotR (Serizay et al., 2020), (Serizay and Ahringer, 2021) was used to analyze fragment distributions and visualized by complexHeatmap (Gu et al., 2016).

### RNA-seq data analysis

All RNA-seq samples were aligned using HISAT2 (Kim et al., 2019) to the mm10 genome. Raw counts were obtained using HT-seq (Putri et al., 2022). TPM values were obtained using Stringtie (Pertea et al., 2015). To compare gene expressions between E6.5 VE RNA-seq (Zhang et al., 2018) and cXEN cells (Fig. 3E), DESeq2's median of ratios (Love et al., 2014) were used for transcript level comparison.

Weighted gene correlation network analysis (WGCNA) (Langfelder and Horvath, 2008) was performed on *Gata6*-FKBP12$^{F36V}$ and Sox17-FKBP12$^{F36V}$ time-course RNA-seq (0, 6, 24, 48, 72, 96h) series using the blockwiseModules function with a "signed" network, TOM type, a soft power threshold of 14, minModuleSize of 150, pamStage set to True, and bicor set to True. Differentially expressed genes for *Fox-A2*-KO cell lines and BMP4 treated samples were analyzed by DESeq2 (Love et al., 2014). RNA-seq heatmaps were visualized using the complexHeatmap package (Gu et al., 2016) or using the Seaborn package (Waskom, 2021).

Gene ontology overrepresentation analyses and GSEA plots were performed using ClusterProfiler (Yu et al., 2012; Wu et al., 2021). GSEA analyses used a rank ordered list from DESeq2 test statistic ("stat" column) output (Love et al., 2014). The GATA6, SOX17, & FOXA2 (GS&F)-cobound VE enhancer targets were generated by extracting differentially expressed genes by DESeq2 by comparing the E6.5 visceral endoderm (VE) versus E6.5 epiblast RNA-seq (Zhang et al., 2018), which were then overlapped with the annotated nearby genes of the GS&F-co-bound VE enhancers (Table S7). The scRNA-seq VE module

set (Table S9) was reduced for the GSEA test by filtering for kME >0.175, and using the top 15% of genes. Statistical analyses were done using the SciPy statistical package using the Wilcox test (Virtanen et al., 2020).

### scRNA-seq data analysis

scRNA-seq data of embryonic day 4.5 (E4.5) embryos were obtained and integrated from Mohammed et al., Cheng et al., Nowotschin et al., Qiu et al. (Mohammed et al., 2017; Cheng et al., 2019; Nowotschin et al., 2019; Qiu et al., 2022) using seurat V4 using standard parameters (Hao et al., 2021). Cell clustering, UMAP, and violinplots for transcriptional signatures were generated by seurat V4 (Hao et al., 2021). The transcriptional signatures were calculated by the UCell algorithm (Hao et al., 2021; Andreatta and Carmona, 2021). Weighted gene correlation network analysis (WGCNA) of single cell transcriptomes were performed by hdWGCNA (Morabito et al., 2021). The parameters used were as follows meta.cell.k = 10, soft_power = 12, minModuleSize = 100, corType = “bicor”. Statistical analyses were done using the Wilcox test through R. To generate cell trajectory (Fig. 1B), we first integrated E4.5 scRNA-seq data (described above) and E5.5 scRNA-seq data from (Nowotschin et al., 2019) using seurat V4. Then, we converted the integrated seurat object to anndata in order to perform analyses of paga, force-directed graph, and pseudotime using scanpy toolkit (Wolf et al., 2018).

### Data and code availability

All raw and processed sequencing data generated in this study have been submitted to the NCBI Gene Expression Omnibus (GEO; https://www.ncbi.nlm.nih.gov/geo/) under accession number GSE213661. The following publicly available datasets were used in this study: E6.5 visceral endoderm RNA-seq, E6.5 epiblast RNA-seq, E6.5 visceral endoderm ATAC-seq, E6.5 visceral endoderm H3K27ac, E6.5 visceral endoderm H3K4me3 can be found at GSE76505 and GSE125318.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgements

We would like to thank Ira Blitz (Cho Lab) for his editorial and critical comments in this work, and Patrick Pham for his assistance with the embryo illustrations in this manuscript. This work was made possible, in part, through access to the Genomics High Throughput Facility Shared Resource of the Cancer Center Support Grant (P30CA-062203) at the University of California, Irvine, NIH shared instrumentation grants 1S10RR025496–01, 1S10OD010794–01, and 1S10OD021718–01, and UC San Diego IGM Genomics Center supported by a National Institutes of Health SIG grant (#S10 OD026929). The authors wish to acknowledge the support of the Chao Family Comprehensive Cancer Center UCI Institute for Immunology Flow Cytometry Facility Shared Resource, supported by the National Cancer Institute of the National Institutes of Health under award number P30CA062203. We thank the University of California, Irvine High Performance Computing Cluster (https://hpc.oit.uci.edu/) for their valuable resources and helpful staff. This research was funded by the following grants awarded to K.W.Y.C.: NIH R35GM139617 and National Science Foundation (NSF) 1755214 and to C.M.: NIH AI00880, AI09599, and AI102853.

## Data availability

Data will be made available on request.

## References


Adamson ED, Ayers SE, 1979. The localization and synthesis of some collagen types in developing mouse embryos. Cell 16, 953–965. 10.1016/0092-8674(79)90110-7. [PubMed: 455457]

Amadei G, Lau KYC, Jonghe JD, Gantner CW, Sozen B, Chan C, Zhu M, Kyprianou C, Hollfelder F, Zernicka-Goetz M, 2021. Inducible stem-cell-derived embryos capture mouse morphogenetic events in vitro. Dev. Cell 56, 366–382. 10.1016/j.devcel.2020.12.004 e9. [PubMed: 33378662]

Alpy F, Jivkov I, Sorokin L, Klein A, Arnold C, Huss Y, Kedinger M, Simon-Assmann P, Lefebvre O, 2005. Generation of a conditionally null allele of the laminin alpha1 gene. Genesis 43, 59–70. [PubMed: 16100707]

Amadei G, Handford CE, Qiu C, Jonghe JD, Greenfeld H, Tran M, Martin BK, Chen D-Y, Aguilera-Castrejon A, Hanna JH, et al. , 2022. Synthetic embryos complete gastrulation to neurulation and organogenesis. Nature 1–3. 10.1038/s41586-022-05246-3.

Andreatta M, Carmona SJ, 2021. UCell: robust and scalable single-cell gene signature scoring. Comput. Struct. Biotechnol. J. 19, 3796–3798. 10.1016/j.csbj.2021.06.043. [PubMed: 34285779]

Ang S-L, Rossant J, 1994. HNF-3β is essential for node and notochord formation in mouse development. Cell 78, 561–574. 10.1016/0092-8674(94)90522-3. [PubMed: 8069909]

Artus J, Panthier J-J, Hadjantonakis A-K, 2010. A role for PDGF signaling in expansion of the extra-embryonic endoderm lineage of the mouse blastocyst. Development 137, 3361–3372. 10.1242/dev.050864. [PubMed: 20826533]

Artus J, Piliszek A, Hadjantonakis A-K, 2011. The primitive endoderm lineage of the mouse blastocyst: sequential transcription factor activation and regulation of differentiation by Sox17. Dev. Biol. 350, 393–404. 10.1016/j.ydbio.2010.12.007. [PubMed: 21146513]

Artus J, Douvaras P, Piliszek A, Isern J, Baron MH, Hadjantonakis A-K, 2012. BMP4 signaling directs primitive endoderm-derived XEN cells to an extraembryonic visceral endoderm identity. Dev. Biol. 361, 245–262. 10.1016/j.ydbio.2011.10.015. [PubMed: 22051107]

Assmat E, Vinot S, Gofflot F, Linsel-Nitschke P, Illien F, Châtelet F, Verroust P, Louvet-Valle S, Rinninger F, Kozyraki R, 2005. Expression and role of cubilin in the internalization of nutrients during the peri-implantation development of the rodent embryo. Biol. Reprod. 72, 1079–1086. 10.1095/biolreprod.104.036913. [PubMed: 15616221]

Barbacci E, Reber M, Ott MO, Breillat C, Huetz F, Cereghini S, 1999. Variant hepatocyte nuclear factor 1 is required for visceral endoderm specification. Development 126, 4795–4805. 10.1242/dev.126.21.4795. [PubMed: 10518496]

Bentsen M, Goymann P, Schultheis H, Klee K, Petrova A, Wiegandt R, Fust A, Preussner J, Kuenne C, Braun T, et al. , 2020. ATAC-seq footprinting unravels kinetics of transcription factor binding during zygotic genome activation. Nat. Commun. 11, 4267. 10.1038/s41467-020-18035-1. [PubMed: 32848148]

Beppu H, Kawabata M, Hamamoto T, Chytil A, Minowa O, Noda T, Miyazono K, 2000. BMP type II receptor is required for gastrulation and early development of mouse embryos. Dev. Biol. 221, 249–258. 10.1006/dbio.2000.9670. [PubMed: 10772805]

Bessonnard S, Mot LD, Gonze D, Barriol M, Dennis C, Goldbeter A, Dupont G, Chazaud C, 2014. Gata6, Nanog and Erk signaling control cell fate in the inner cell mass through a tristable regulatory network. Development 141, 3637–3648. 10.1242/dev.109678. [PubMed: 25209243]

Bolger AM, Lohse M, Usadel B, 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30, 2114–2120. 10.1093/bioinformatics/btu170. [PubMed: 24695404]

Bowles J, Schepers G, Koopman P, 2000. Phylogeny of the SOX family of developmental transcription factors based on sequence and structural indicators. Dev. Biol. 227, 239–255. 10.1006/dbio.2000.9883. [PubMed: 11071752]

Burtscher I, Lickert H, 2009. Foxa2 regulates polarity and epithelialization in the endoderm germ layer of the mouse embryo. Development 136, 1029–1038. 10.1242/dev.028415. [PubMed: 19234065]

Cai KQ, Capo-Chichi CD, Rula ME, Yang D, Xu X, 2008. Dynamic GATA6 expression in primitive endoderm formation and maturation in early mouse embryogenesis. Dev. Dynam. 237, 2820–2829. 10.1002/dvdy.21703.

Casanova JE, Grabel LB, 1988. The role of cell interactions in the differentiation of teratocarcinoma-derived parietal and visceral endoderm. Dev. Biol. 129, 124–139. 10.1016/0012-1606(88)90167-4. [PubMed: 2457525]

Castro-Mondragon JA, Riudavets-Puig R, Rauluseviciute I, Berhanu Lemma R, Turchi L, Blanc-Mathieu R, Lucas J, Boddie P, Khan A, Manosalva Pérez N, et al. , 2021. JASPAR 2022: the 9th release of the open-access database of transcription factor binding profiles. Nucleic Acids Res. 50, gkab1113. 10.1093/nar/gkab1113.

Chazaud C, Yamanaka Y, Pawson T, Rossant J, 2006. Early lineage segregation between epiblast and primitive endoderm in mouse blastocysts through the grb2-MAPK pathway. Dev. Cell 10, 615–624. 10.1016/j.devcel.2006.02.020. [PubMed: 16678776]

Chen WS, Manova K, Weinstein DC, Duncan SA, Plump AS, Prezioso VR, Bachvarova RF, Darnell JE, 1994. Disruption of the HNF-4 gene, expressed in visceral endoderm, leads to cell death in embryonic ectoderm and impaired gastrulation of mouse embryos. Gene Dev. 8, 2466–2477. 10.1101/gad.8.20.2466. [PubMed: 7958910]

Cheng S, Pei Y, He L, Peng G, Reinius B, Tam PPL, Jing N, Deng Q, 2019. Single-cell RNA-seq reveals cellular heterogeneity of pluripotency transition and X chromosome dynamics during early mouse development. Cell Rep. 26, 2593–2607. 10.1016/j.celrep.2019.02.031 e3. [PubMed: 30840884]

Cho LTY, Wamaitha SE, Tsai IJ, Artus J, Sherwood RI, Pedersen RA, Hadjantonakis A-K, Niakan KK, 2012. Conversion from mouse embryonic to extra-embryonic endoderm stem cells reveals distinct differentiation capacities of pluripotent stem cell states. Development 139, 2866–2877. 10.1242/dev.078519. [PubMed: 22791892]

Chu GC, Dunn NR, Anderson DC, Oxburgh L, Robertson EJ, 2004. Differential requirements for Smad4 in TGFβ-dependent patterning of the early mouse embryo. Development 131, 3501–3512. 10.1242/dev.01248. [PubMed: 15215210]

Cirillo LA, Lin FR, Cuesta I, Friedman D, Jarnik M, Zaret KS, 2002. Opening of compacted chromatin by early developmental transcription factors HNF3 (FoxA) and GATA-4. Mol. Cell. 9, 279–289. 10.1016/s1097-2765(02)00459-8. [PubMed: 11864602]

Cockroft DL, Gardner RL, 1987. Clonal analysis of the developmental potential of 6th and 7th day visceral endoderm cells in the mouse. Development 101, 143–155. 10.1242/dev.101.1.143. [PubMed: 3503712]

Coffinier C, Thepot D, Babinet C, Yaniv M, Barra J, 1999. Essential role for the homeoprotein vHNF1/HNF1beta in visceral endoderm differentiation. Development 126, 4785–4794. 10.1242/dev.126.21.4785. [PubMed: 10518495]

Cooper AR, Kurkinen M, Taylor A, Hogan BLM, 1981. Studies on the biosynthesis of laminin by murine parietal endoderm cells. Eur. J. Biochem. 119, 189–197. 10.1111/j.1432-1033.1981.tb05593.x. [PubMed: 7341241]

Costello I, Nowotschin S, Sun X, Mould AW, Hadjantonakis A-K, Bikoff EK, Robertson EJ, 2015. Lhx1 functions together with Otx2, Foxa2, and Ldb1 to govern anterior mesendoderm, node, and midline development. Gene Dev. 29, 2108–2122. 10.1101/gad.268979.115. [PubMed: 26494787]

Coucouvanis E, Martin GR, 1999. BMP signaling plays a role in visceral endoderm differentiation and cavitation in the early mouse embryo. Development 126, 535–546. 10.1242/dev.126.3.535. [PubMed: 9876182]

Derynck R, Budi EH, 2019. Specificity, versatility, and control of TGF-β family signaling. Sci. Signal. 12, eaav5183. 10.1126/scisignal.aav5183. [PubMed: 30808818]

Dietrich J-E, Hiiragi T, 2007. Stochastic patterning in the mouse pre-implantation embryo. Development 134, 4219–4231. 10.1242/dev.003798. [PubMed: 17978007]

Donovan A, Lima CA, Pinkus JL, Pinkus GS, Zon LI, Robine S, Andrews NC, 2005. The iron exporter ferroportin/Slc40a1 is essential for iron homeostasis. Cell Metabol. 1, 191–200. 10.1016/j.cmet.2005.01.003.

Dufort D, Schwartz L, Harpal K, Rossant J, 1998. The transcription factor HNF3beta is required in visceral endoderm for normal primitive streak morphogenesis. Development 125, 3015–3025. 10.1242/dev.125.16.3015. [PubMed: 9671576]

Duncan SA, Nagy A, Chan W, 1997. Murine gastrulation requires HNF-4 regulated gene expression in the visceral endoderm: tetraploid rescue of Hnf-4(−/−) embryos. Development 124, 279–287. 10.1242/dev.124.2.279. [PubMed: 9053305]

Ferner K, Mess A, 2011. Evolution and development of fetal membranes and placentation in amniote vertebrates. Respir. Physiol. Neurobiol. 178, 39–50. 10.1016/j.resp.2011.03.029. [PubMed: 21470579]

Filimonow K, Fuente R. de la, 2021. Specification and role of extraembryonic endoderm lineages in the peri-implantation mouse embryo. Theriogenology 180, 189–206. 10.1016/j.theriogenology.2021.12.021. [PubMed: 34998083]

Fowler KJ, Mitrangas K, Dziadek M, 1990. In vitro production of Reichert's membrane by mouse embryo-derived parietal endoderm cell lines. Exp. Cell Res. 191, 194–203. 10.1016/0014-4827(90)90005-u. [PubMed: 2124186]

Fujiwara Y, Tanno Y, Sugishita H, Kishi Y, Makino Y, Okada Y, 2021. Preparation of optimized concanavalin A-conjugated Dynabeads® magnetic beads for CUT&Tag. PLoS One 16, e0259846. 10.1371/journal.pone.0259846. [PubMed: 34784358]

Graham SJL, Wicher KB, Jedrusik A, Guo G, Herath W, Robson P, Zernicka-Goetz M, 2014. BMP signalling regulates the pre-implantation development of extraembryonic cell lineages in the mouse embryo. Nat. Commun. 5, 5667. 10.1038/ncomms6667. [PubMed: 25514175]

Gu Z, Eils R, Schlesner M, 2016. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. Bioinformatics 32, 2847–2849. 10.1093/bioinformatics/btw313. [PubMed: 27207943]

Hamilton WB, Mosesson Y, Monteiro RS, Emdal KB, Knudsen TE, Francavilla C, Barkai N, Olsen JV, Brickman JM, 2019. Dynamic lineage priming is driven via direct enhancer regulation by ERK. Nature 575, 355–360. 10.1038/s41586-019-1732-z. [PubMed: 31695196]

Haniel A, Welge-Lüssen U, Kühn K, Pöschl E, 1995. Identification and characterization of a novel transcriptional silencer in the human collagen type IV gene COL4A2. J. Biol. Chem. 270 (19), 11209–11215. [PubMed: 7744753]

Hao Y, Hao S, Andersen-Nissen E, Mauck WM, Zheng S, Butler A, Lee MJ, Wilk AJ, Darby C, Zager M, et al. , 2021. Integrated analysis of multimodal single-cell data. Cell 184, 3573–3587. 10.1016/j.cell.2021.04.048 e29. [PubMed: 34062119]

Hayamizu TF, Wicks MN, Davidson DR, Burger A, Ringwald M, Baldock RA, 2013. EMAP/EMAPA ontology of mouse developmental anatomy: 2013 update. J. Biomed. Semant. 4, 15. 10.1186/2041-1480-4-15.

Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, Cheng JX, Murre C, Singh H, Glass CK, 2010. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol. Cell. 38, 576–589. 10.1016/j.molcel.2010.05.004. [PubMed: 20513432]

Holmes WR, Mochel N.S.R. de, Wang Q, Du H, Peng T, Chiang M, Cinquin O, Cho K, Nie Q, 2017. Gene expression noise enhances robust organization of the early mammalian blastocyst. PLoS Comput. Biol. 13, e1005320. 10.1371/journal.pcbi.1005320. [PubMed: 28114387]

Javier AL, Doan LT, Luong M, Mochel N.S.R. de, Sun A, Monuki ES, Cho KWY, 2012. Bmp indicator mice reveal dynamic regulation of transcriptional response. PLoS One 7, e42566. 10.1371/journal.pone.0042566. [PubMed: 22984405]

Jindal GA, Farley EK, 2021. Enhancer grammar in development, evolution, and disease: dependencies and interplay. Dev. Cell 56, 575–587. 10.1016/j.devcel.2021.02.016. [PubMed: 33689769]

Kanai-Azuma M, Kanai Y, Gad JM, Tajima Y, Taya C, Kurohmaru M, Sanai Y, Yonekawa H, Yazaki K, Tam PPL, et al. , 2002. Depletion of definitive gut endoderm in Sox17 -null mutant mice. Development 129, 2367–2379. 10.1242/dev.129.10.2367. [PubMed: 11973269]

Kanehisa M, Goto S, 2000. KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res. 28, 27–30. 10.1093/nar/28.1.27. [PubMed: 10592173]

Kang M, Piliszek A, Artus J, Hadjantonakis A-K, 2013. FGF4 is required for lineage restriction and salt-and-pepper distribution of primitive endoderm factors but not their initial expression in the mouse. Development 140, 267–279. 10.1242/dev.043471. [PubMed: 23193166]

Kang M, Garg V, Hadjantonakis A-K, 2017. Lineage establishment and progression within the inner cell mass of the mouse blastocyst requires FGFR1 and FGFR2. Dev. Cell 41, 496–510. 10.1016/j.devcel.2017.05.003 e5. [PubMed: 28552559]

Kawamura N, Sun-Wada G-H, Aoyama M, Harada A, Takasuga S, Sasaki T, Wada Y, 2012. Delivery of endosomes to lysosomes via microautophagy in the visceral endoderm of mouse embryos. Nat. Commun. 3, 1071. 10.1038/ncomms2069. [PubMed: 22990867]

Kawamura N, Takaoka K, Hamada H, Hadjantonakis A-K, Sun-Wada G-H, Wada Y, 2020. Rab7-Mediated endocytosis establishes patterning of wnt activity through inactivation of dkk antagonism. Cell Rep. 31, 107733. 10.1016/j.celrep.2020.107733. [PubMed: 32521258]

Kim D, Paggi JM, Park C, Bennett C, Salzberg SL, 2019. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. Nat. Biotechnol. 37, 907–915. 10.1038/s41587-019-0201-4. [PubMed: 31375807]

Koutsourakis M, Langeveld A, Patient R, Beddington R, Grosveld F, 1999. The transcription factor GATA6 is essential for early extraembryonic development. Development 126, 723–732. 10.1242/dev.126.9.723.

Kulakovskiy IV, Vorontsov IE, Yevshin IS, Sharipov RN, Fedorova AD, Rumynskiy EI, Medvedeva YA, Magana-Mora A, Bajic VB, Papatsenko DA, et al. , 2017. HOCOMOCO: towards a complete collection of transcription factor binding models for human and mouse via large-scale ChIP-Seq analysis. Nucleic Acids Res. 46, gkx1106. 10.1093/nar/gkx1106.

Kunath T, Arnaud D, Uy GD, Okamoto I, Chureau C, Yamanaka Y, Heard E, Gardner RL, Avner P, Rossant J, 2005. Imprinted X-inactivation in extraembryonic endoderm cell lines from mouse blastocysts. Development 132, 1649–1661. 10.1242/dev.01715. [PubMed: 15753215]

Kuo CT, Morrisey EE, Anandappa R, Sigrist K, Lu MM, Parmacek MS, Soudais C, Leiden JM, 1997. GATA4 transcription factor is required for ventral morphogenesis and heart tube formation. Gene Dev. 11, 1048–1060. 10.1101/gad.11.8.1048. [PubMed: 9136932]

Kwon GS, Hadjantonakis A, 2009. Transthyretin mouse transgenes direct RFP expression or Cre-mediated recombination throughout the visceral endoderm. Genesis 47, 447–455. 10.1002/dvg.20522. [PubMed: 19415627]

Lane EB, Hogan BLM, Kurkinen M, Garrels JI, 1983. Co-expression of vimentin and cytokeratins in parietal endoderm cells of early mouse embryo. Nature 303, 701–704. 10.1038/303701a0. [PubMed: 6190091]

Langfelder P, Horvath S, 2008. WGCNA: an R package for weighted correlation network analysis. BMC Bioinf. 9, 559. 10.1186/1471-2105-9-559.

Langmead B, Salzberg SL, 2012. Fast gapped-read alignment with Bowtie 2. Nat. Methods 9, 357–359. 10.1038/nmeth.1923. [PubMed: 22388286]

Leung CY, Zernicka-Goetz M, 2013. Angiomotin prevents pluripotent lineage differentiation in mouse embryos via Hippo pathway-dependent and -independent mechanisms. Nat. Commun. 4, 2251. 10.1038/ncomms3251. [PubMed: 23903990]

Lim CY, Tam W-L, Zhang J, Ang HS, Jia H, Lipovich L, Ng H-H, Wei C-L, Sung WK, Robson P, et al. , 2008. Sall4 regulates distinct transcription circuitries in different blastocyst-derived stem cell lineages. Cell Stem Cell 3, 543–554. 10.1016/j.stem.2008.08.004. [PubMed: 18804426]

Liu N, 2019. Library Prep for CUT&RUN with NEBNext® Ultra™ II DNA Library Prep Kit for Illumina® (E7645) v2. 10.17504/protocols.io.bagaibse.

Liu N, Hargreaves VV, Zhu Q, Kurland JV, Hong J, Kim W, Sher F, Macias-Trevino C, Rogers JM, Kurita R, et al. , 2018. Direct promoter repression by BCL11A controls the fetal to adult hemoglobin switch. Cell 173, 430–442. 10.1016/j.cell.2018.03.016 e17. [PubMed: 29606353]

Love MI, Huber W, Anders S, 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15, 550. 10.1186/s13059-014-0550-8. [PubMed: 25516281]

McDonald ACH, Biechele S, Rossant J, Stanford WL, 2014. Sox17-Mediated XEN cell conversion identifies dynamic networks controlling cell-fate decisions in embryo-derived stem cells. Cell Rep. 9, 780–793. 10.1016/j.celrep.2014.09.026. [PubMed: 25373912]

Meehan RR, Barlow DP, Hill RE, Hogan BL, Hastie ND, 1984. Pattern of serum protein gene expression in mouse visceral yolk sac and foetal liver. EMBO J. 3, 1881–1885. 10.1002/j.1460-2075.1984.tb02062.x. [PubMed: 6479150]

Meers MP, Bryson TD, Henikoff JG, Henikoff S, 2019. Improved CUT&RUN chromatin profiling tools. Elife 8, e46314. 10.7554/elife.46314. [PubMed: 31232687]

Miner JH, Li C, Mudd JL, Go G, Sutherland AE, 2004. Compositional and structural requirements for laminin and basement membranes during mouse embryo implantation and gastrulation. Development 131, 2247–2256. [PubMed: 15102706]

Mochel N.S.R. de, Luong M, Chiang M, Javier AL, Luu E, Toshihiko F, MacGregor GR, Cinquin O, Cho KWY, 2015. BMP signaling is required for cell cleavage in preimplantation-mouse embryos. Dev. Biol. 397, 45–55. 10.1016/j.ydbio.2014.10.001. [PubMed: 25446538]

Mohammed H, Hernando-Herraez I, Savino A, Scialdone A, Macaulay I, Mulas C, Chandra T, Voet T, Dean W, Nichols J, et al. , 2017. Single-cell landscape of transcriptional heterogeneity and cell fate decisions during mouse early gastrulation. Cell Rep. 20, 1215–1228. 10.1016/j.celrep.2017.07.009. [PubMed: 28768204]

Molkentin JD, Lin Q, Duncan SA, Olson EN, 1997. Requirement of the transcription factor GATA4 for heart tube formation and ventral morphogenesis. Gene Dev. 11, 1061–1072. 10.1101/gad.11.8.1061. [PubMed: 9136933]

Molotkov A, Mazot P, Brewer JR, Cinalli RM, Soriano P, 2017. Distinct requirements for FGFR1 and FGFR2 in primitive endoderm development and exit from pluripotency. Dev. Cell 41, 511–526. 10.1016/j.devcel.2017.05.004 e4. [PubMed: 28552557]

Morabito S, Miyoshi E, Michael N, Shahin S, Martini AC, Head E, Silva J, Leavy K, Perez-Rosendahl M, Swarup V, 2021. Single-nucleus chromatin accessibility and transcriptomic characterization of Alzheimer's disease. Nat. Genet. 53, 1143–1155. 10.1038/s41588-021-00894-z. [PubMed: 34239132]

Morris SA, Teo RTY, Li H, Robson P, Glover DM, Zernicka-Goetz M, 2010. Origin and formation of the first two distinct cell types of the inner cell mass in the mouse embryo. Proc. Natl. Acad. Sci. USA 107, 6364–6369. 10.1073/pnas.0915063107. [PubMed: 20308546]

Nabet B, Roberts JM, Buckley DL, Paulk J, Dastjerdi S, Yang A, Leggett AL, Erb MA, Lawlor MA, Souza A, et al. , 2018. The dTAG system for immediate and target-specific protein degradation. Nat. Chem. Biol. 14, 431–441. 10.1038/s41589-018-0021-8. [PubMed: 29581585]

Nabet B, Ferguson FM, Seong BKA, Kuljanin M, Leggett AL, Mohardt ML, Robichaud A, Conway AS, Buckley DL, Mancias JD, et al. , 2020. Rapid and direct control of target protein levels with VHL-recruiting dTAG molecules. Nat. Commun. 11, 4687. 10.1038/s41467-020-18377-w. [PubMed: 32948771]

Niakan KK, Ji H, Maehr R, Vokes SA, Rodolfa KT, Sherwood RI, Yamaki M, Dimos JT, Chen AE, Melton DA, et al. , 2010. Sox17 promotes differentiation in mouse embryonic stem cells by directly regulating extraembryonic gene expression and indirectly antagonizing self-renewal. Gene Dev. 24, 312–326. 10.1101/gad.1833510. [PubMed: 20123909]

Niakan KK, Schrode N, Cho LTY, Hadjantonakis A-K, 2013. Derivation of extraembryonic endoderm stem (XEN) cells from mouse embryos and embryonic stem cells. Nat. Protoc. 8, 1028–1041. 10.1038/nprot.2013.049. [PubMed: 23640167]

Niimi T, Hayashi Y, Futaki S, Sekiguchi K, 2004. SOX7 and SOX17 regulate the parietal endoderm-specific enhancer activity of mouse laminin α1 gene. J. Biol. Chem. 279, 38055–38061. 10.1074/jbc.m403724200. [PubMed: 15220343]

Ninomiya Y, Davies TJ, Gardner RL, 2005. Experimental analysis of the transdifferentiation of visceral to parietal endoderm in the mouse. Dev. Dynam. 233, 837–846. 10.1002/dvdy.20405.

Nishioka N, Inoue K, Adachi K, Kiyonari H, Ota M, Ralston A, Yabuta N, Hirahara S, Stephenson RO, Ogonuki N, et al. , 2009. The Hippo signaling pathway components lats and yap pattern Tead4 activity to distinguish mouse trophectoderm from inner cell mass. Dev. Cell 16, 398–410. 10.1016/j.devcel.2009.02.003. [PubMed: 19289085]

Niwa H, Toyooka Y, Shimosato D, Strumpf D, Takahashi K, Yagi R, Rossant J, 2005. Interaction between oct3/4 and Cdx2 determines trophectoderm differentiation. Cell 123, 917–929. 10.1016/j.cell.2005.08.040. [PubMed: 16325584]

Nowotschin S, Setty M, Kuo Y-Y, Liu V, Garg V, Sharma R, Simon CS, Saiz N, Gardner R, Boutet SC, et al. , 2019. The emergent landscape of the mouse gut endoderm at single-cell resolution. Nature 569, 361–367. 10.1038/s41586-019-1127-1. [PubMed: 30959515]

Ohinata Y, Endo TA, Sugishita H, Watanabe T, Iizuka Y, Kawamoto Y, Saraya A, Kumon M, Koseki Y, Kondo T, et al. , 2022. Establishment of mouse stem cells that can recapitulate the developmental potential of primitive endoderm. Science 375, 574–578. 10.1126/science.aay3325. [PubMed: 35113719]

Paca A, Seguin CA, Clements M, Ryczko M, Rossant J, Rodriguez TA, Kunath T, 2012. BMP signaling induces visceral endoderm differentiation of XEN cells and parietal endoderm. Dev. Biol. 361, 90–102. 10.1016/j.ydbio.2011.10.013. [PubMed: 22027433]

Perea-Gomez A, Cases O, Lelièvre V, Pulina MV, Collignon J, Hadjantonakis A-K, Kozyraki R, 2019. Loss of Cubilin, the intrinsic factor-vitamin B12 receptor, impairs visceral endoderm endocytosis and endodermal patterning in the mouse. Sci Rep-uk 9, 10168. 10.1038/s41598-019-46559-0.

Pertea M, Pertea GM, Antonescu CM, Chang T-C, Mendell JT, Salzberg SL, 2015. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. Nat. Biotechnol. 33, 290–295. 10.1038/nbt.3122. [PubMed: 25690850]

Plusa B, Piliszek A, Frankenberg S, Artus J, Hadjantonakis A-K, 2008. Distinct sequential cell behaviours direct primitive endoderm formation in the mouse blastocyst. Development 135, 3081–3091. 10.1242/dev.021519. [PubMed: 18725515]

Putri GH, Anders S, Pyl PT, Pimanda JE, Zanini F, 2022. Analysing high-throughput sequencing data in Python with HTSeq 2.0. Bioinformatics 38, 2943–2945. 10.1093/bioinformatics/btac166. [PubMed: 35561197]

Qiu C, Cao J, Martin BK, Li T, Welsh IC, Srivatsan S, Huang X, Calderon D, Noble WS, Disteche CM, et al. , 2022. Systematic reconstruction of cellular trajectories across mouse embryogenesis. Nat. Genet. 54, 328–341. 10.1038/s41588-022-01018-x. [PubMed: 35288709]

Qu X, Pan J, Zhang C, Huang S, 2008. Sox17 facilitates the differentiation of mouse embryonic stem cells into primitive and definitive endoderm in vitro. Dev. Growth Differ. 50, 585–593. 10.1111/j.1440-169x.2008.01056.x. [PubMed: 19238729]

Ramírez F, Dündar F, Diehl S, Grüning BA, Manke T, 2014. deepTools: a flexible platform for exploring deep-sequencing data. Nucleic Acids Res. 42, W187–W191. 10.1093/nar/gku365. [PubMed: 24799436]

Rayon T, Menchero S, Nieto A, Xenopoulos P, Crespo M, Cockburn K, Cañon S, Sasaki H, Hadjantonakis A-K, de la Pompa JL, et al. , 2014. Notch and Hippo converge on Cdx2 to specify the trophectoderm lineage in the mouse blastocyst. Dev. Cell 30, 410–422. 10.1016/j.devcel.2014.06.019. [PubMed: 25127056]

Sahoo SS, Ramanand SG, Gao Y, Abbas A, Kumar A, Cuevas IC, Li H-D, Aguilar M, Xing C, Mani RS, et al. , 2022. FOXA2 suppresses endometrial carcinogenesis and epithelial-mesenchymal transition by regulating enhancer activity. J. Clin. Investig. 132, e157574. 10.1172/jci157574. [PubMed: 35703180]

Saiz N, Williams KM, Seshan VE, Hadjantonakis A-K, 2016. Asynchronous fate decisions by single cells collectively ensure consistent lineage composition in the mouse blastocyst. Nat. Commun. 7, 13463. 10.1038/ncomms13463. [PubMed: 27857135]

Salamat M, Götz W, Horster A, Janotte B, Herken R, 1993. Ultrastructural localization of carbohydrates in Reichert's membrane of the mouse. Cell Tissue Res. 272, 375–381. 10.1007/BF00302742. [PubMed: 8513488]

Salamat M, Miosge N, Herken R, 1995. Development of Reichert's membrane in the early mouse embryo. Anat. Embryol. 192, 275–281. 10.1007/BF00184752.

Schrode N, Saiz N, Talia SD, Hadjantonakis A-K, 2014. GATA6 levels modulate primitive endoderm cell fate choice and timing in the mouse blastocyst. Dev. Cell 29, 454–467. 10.1016/j.devcel.2014.04.011. [PubMed: 24835466]

Semoff S, Hogan BL, Hopkins CR, 1982. Localization of fibronectin, laminin-entactin, and entactin in Reichert's membrane by immunoelectron microscopy. EMBO J. 1, 1171–1175. 10.1002/j.1460-2075.1982.tb00009.x. [PubMed: 7188248]

Serizay J, Ahringer J, 2021. Generating fragment density plots in R/Bioconductor with VplotR. J. Open Source Softw. 6, 3009. 10.21105/joss.03009.

Serizay J, Dong Y, Janes J, Chesney M, Cerrato C, Ahringer J, 2020. Distinctive regulatory architectures of germline-active and somatic genes in C. elegans. Genome Res. 30. 10.1101/gr.265934.120gr.265934.120.

Sheng G, Foley AC, 2012. Diversification and conservation of the extraembryonic tissues in mediating nutrient uptake during amniote development. Ann Ny Acad Sci 1271, 97–103. 10.1111/j.1749-6632.2012.06726.x. [PubMed: 23050970]

Shimoda M, Kanai-Azuma M, Hara K, Miyazaki S, Kanai Y, Monden M, Miyazaki J, 2007. Sox17 plays a substantial role in late-stage differentiation of the extraembryonic endoderm in vitro. J. Cell Sci. 120, 3859–3869. 10.1242/jcs.007856. [PubMed: 17940068]

Sirard C, de la Pompa JL, Elia A, Itie A, Mirtsos C, Cheung A, Hahn S, Wakeham A, Schwartz L, Kern SE, et al. , 1998. The tumor suppressor gene Smad4/Dpc4 is required for gastrulation and later for anterior development of the mouse embryo. Gene Dev. 12, 107–119. 10.1101/gad.12.1.107. [PubMed: 9420335]

Skene PJ, Henikoff S, 2017. An efficient targeted nuclease strategy for high-resolution mapping of DNA binding sites. Elife 6, e21856. 10.7554/elife.21856. [PubMed: 28079019]

Smyth N, Vatansever HS, Murray P, Meyer M, Frie C, Paulsson M, Edgar D, 1999. Absence of basement membranes after targeting the LAMC1 gene results in embryonic lethality due to failure of endoderm differentiation. J. Cell bio. 144 (1), 151–160. [PubMed: 9885251]

Sozen B, Amadei G, Cox A, Wang R, Na E, Czukiewska S, Chappell L, Voet T, Michel G, Jing N, et al. , 2018. Self-assembly of embryonic and two extraembryonic stem cell types into gastrulating embryo-like structures. Nat. Cell Biol. 20, 979–989. 10.1038/s41556-018-0147-7. [PubMed: 30038254]

Sozen B, Demir N, Zernicka-Goetz M, 2021. BMP signalling is required for extraembryonic ectoderm development during pre-to-post-implantation transition of the mouse embryo. Dev. Biol. 470, 84–94. 10.1016/j.ydbio.2020.11.005. [PubMed: 33217407]

Strickland S, Mahdavi V, 1978. The induction of differentiation in teratocarcinoma stem cells by retinoic acid. Cell 15, 393–403. 10.1016/0092-8674(78)90008-9. [PubMed: 214238]

Strickland S, Smith KK, Marotti KR, 1980. Hormonal induction of differentiation in teratocarcinoma stem cells: generation of parietal endoderm by retinoic acid and dibutyryl cAMP. Cell 21, 347–355. 10.1016/0092-8674(80)90471-7. [PubMed: 6250719]

Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, et al. , 2005. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc. Natl. Acad. Sci. USA 102, 15545–15550. 10.1073/pnas.0506580102. [PubMed: 16199517]

Sun-Wada G-H, Wada Y, 2022. The a subunit isoforms of vacuolar-type proton ATPase exhibit differential distribution in mouse perigastrulation embryos. Sci Rep-uk 12, 13590. 10.1038/s41598-022-18002-4.

Sun-Wada G-H, Tabata H, Wada Y, 2021. Vacuolar-type proton ATPase is required for maintenance of apicobasal polarity of embryonic visceral endoderm. Sci Rep-uk 11, 19355. 10.1038/s41598-021-98952-3.

Tarazi S, Aguilera-Castrejon A, Joubran C, Ghanem N, Ashouokhi S, Roncato F, Wildschutz E, Haddad M, Oldak B, Gomez-Cesar E, et al. , 2022. Post-gastrulation synthetic embryos generated ex utero from mouse naive ESCs. Cell 185, 3290–3306.e25. 10.1016/j.cell.2022.07.028. [PubMed: 35988542]

Ueda Y, Kimura-Yoshida C, Mochida K, Tsume M, Kameo Y, Adachi T, Lefebvre O, Hiramatsu R, Matsuo I, 2020. Intrauterine pressures adjusted by Reichert's membrane are crucial for early mouse morphogenesis. Cell Rep. 31, 107637. 10.1016/j.celrep.2020.107637. [PubMed: 32433954]

van de Stolpe A, Karperien M, Löwik CW, Jüppner H, Segre GV, Abou-Samra AB, de Laat SW, Defize LH, 1993. Parathyroid hormone-related peptide as an endogenous inducer of parietal endoderm differentiation. J. Cell Biol. 120, 235–243. 10.1083/jcb.120.1.235. [PubMed: 8380175]

Verheijen MHG, Karperien M, Chung U, Wijuen M. van, Heystek H, Hendriks JAA, Veltmaat JM, Lanske B, Li E, Löwik CWGM, et al. , 1999. Parathyroid hormone-related peptide (PTHrP)

induces parietal endoderm formation exclusively via the Type I PTH/PTHrP receptor. Mech. Dev. 81, 151–161. 10.1016/s0925-4773(98)00240-8. [PubMed: 10330492]

Virtanen P, Gommers R, Oliphant TE, Haberland M, Reddy T, Cournapeau D, Burovski E, Peterson P, Weckesser W, Bright J, et al. , 2020. SciPy 1.0: fundamental algorithms for scientific computing in Python. Nat. Methods 17, 261–272. 10.1038/s41592-019-0686-2. [PubMed: 32015543]

Wallingford MC, Angelo JR, Mager J, 2013. Morphogenetic analysis of periimplantation development. Dev. Dynam. 242, 1110–1120. 10.1002/dvdy.23991.

Wamaitha SE, Valle I. del, Cho LTY, Wei Y, Fogarty NME, Blakeley P, Sherwood RI, Ji H, Niakan KK, 2015. Gata6 potently initiates reprograming of pluripotent and differentiated cells to extraembryonic endoderm stem cells. Gene Dev. 29, 1239–1255. 10.1101/gad.257071.114. [PubMed: 26109048]

Waskom M, 2021. seaborn: statistical data visualization. J. Open Source Softw. 6, 3021. 10.21105/joss.03021.

Wat MJ, Beck TF, Hernandez-García A, Yu Z, Veenma D, Garcia M, Holder AM, Wat JJ, Chen Y, Mohila CA, et al. , 2012. Mouse model reveals the role of SOX7 in the development of congenital diaphragmatic hernia associated with recurrent deletions of 8p23.1. Hum. Mol. Genet. 21, 4115–4125. 10.1093/hmg/dds241. [PubMed: 22723016]

Weinstein DC, Altaba A.R. i, Chen WS, Hoodless P, Prezioso VR, Jessell TM, Darnell JE, 1994. The winged-helix transcription factor HNF-3β is required for notochord development in the mouse embryo. Cell 78, 575–588. 10.1016/0092-8674(94)90523-1. [PubMed: 8069910]

Wolf FA, Angerer P, Theis FJ, 2018. SCANPY: large-scale single-cell gene expression data analysis. Genome Biol. 19, 15. 10.1186/s13059-017-1382-0. [PubMed: 29409532]

Wu T, Hu E, Xu S, Chen M, Guo P, Dai Z, Feng T, Zhou L, Tang W, Zhan L, et al. , 2021. clusterProfiler 4.0: a universal enrichment tool for interpreting omics data. Innovation 2, 100141. 10.1016/j.xinn.2021.100141. [PubMed: 34557778]

Xenopoulos P, Kang M, Puliafito A, Di Talia S, Hadjantonakis A-K, 2015. Heterogeneities in nanog expression drive stable commitment to pluripotency in the mouse blastocyst. Cell Rep. 10, 1508–1520. 10.1016/j.celrep.2015.02.010. [PubMed: 25753417]

Xiang Y, Zhang Y, Xu Q, Zhou C, Liu B, Du Z, Zhang K, Zhang B, Wang X, Gayen S, et al. , 2020. Epigenomic analysis of gastrulation identifies a unique chromatin state for primed pluripotency. Nat. Genet. 52, 95–105. 10.1038/s41588-019-0545-1. [PubMed: 31844322]

Yagi S, Tagawa Y, Shiojiri N, 2016. Transdifferentiation of mouse visceral yolk sac cells into parietal yolk sac cells in vitro. Biochem Bioph Res Co 470, 917–923. 10.1016/j.bbrc.2016.01.149.

Yamanaka Y, Lanner F, Rossant J, 2010. FGF signal-dependent segregation of primitive endoderm and epiblast in the mouse blastocyst. Development 137, 715–724. 10.1242/dev.043471. [PubMed: 20147376]

Yang X, Li C, Xu X, Deng C, 1998. The tumor suppressor SMAD4/DPC4 is essential for epiblast proliferation and mesoderm induction in mice. Proc. Natl. Acad. Sci. USA 95, 3667–3672. 10.1073/pnas.95.7.3667. [PubMed: 9520423]

Yu G, Wang L-G, Han Y, He Q-Y, 2012. clusterProfiler: an R Package for comparing biological themes among gene clusters. OMICS A J. Integr. Biol. 16, 284–287. 10.1089/omi.2011.0118.

Yu S, Yehia G, Wang J, Stypulkowski E, Sakamori R, Jiang P, Hernandez-Enriquez B, Tran TS, Bonder EM, Guo W, et al. , 2014. Global ablation of the mouse Rab11a gene impairs early embryogenesis and matrix metalloproteinase secretion. J. Biol. Chem. 289, 32030–32043. 10.1074/jbc.M113.538223. [PubMed: 25271168]

Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, et al. , 2008. Model-based analysis of ChIP-seq (MACS). Genome Biol. 9, R137. 10.1186/gb-2008-9-9-r137. [PubMed: 18798982]

Zhang Y, Xiang Y, Yin Q, Du Z, Peng X, Wang Q, Fidalgo M, Xia W, Li Y, Zhao Z, et al. , 2018. Dynamic epigenomic landscapes during early lineage specification in mouse embryos. Nat. Genet. 50, 96–105. 10.1038/s41588-017-0003-x. [PubMed: 29203909]

Zhang S, Chen T, Chen N, Gao D, Shi B, Kong S, West RC, Yuan Y, Zhi M, Wei Q, et al. , 2019. Implantation initiation of self-assembled embryo-like structures generated using three types

of mouse blastocyst-derived stem cells. Nat. Commun. 10, 496. 10.1038/s41467-019-08378-9. [PubMed: 30700702]

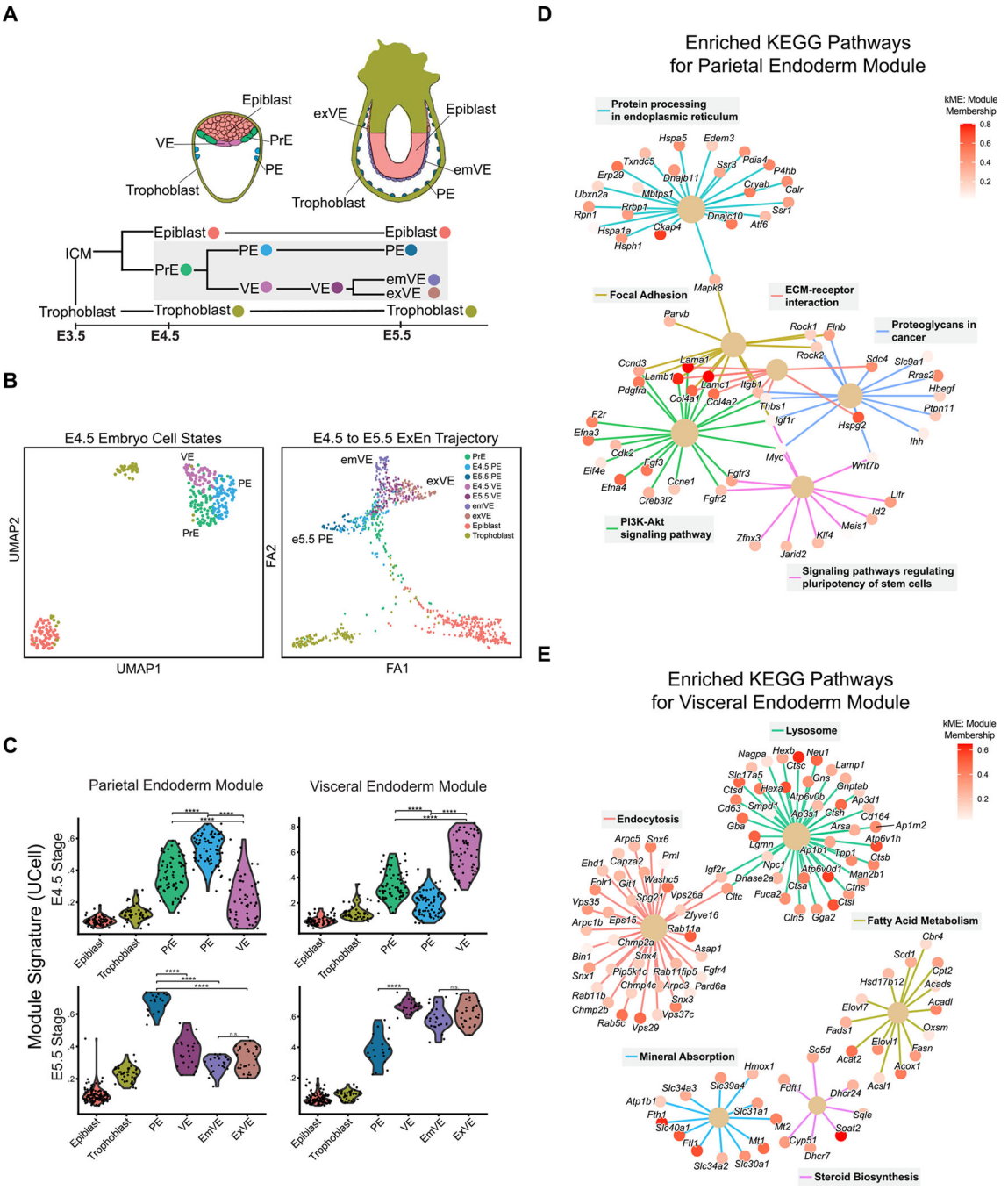


**Fig. 1. Single-cell transcriptomes of the extraembryonic endoderm lineages in E4.5 embryo.**

(A) A schematic highlighting embryonic stages from E3.5 and E5.5, depicting the lineage relationships of cell states with a primary emphasis on PrE lineage segregation events from E4.5 to E5.5. (B) Left panel displays a UMAP embedding of the integrated scRNA-seq data of the E4.5 embryo from Nowotschin et al. (Nowotschin et al., 2019) and Qiu et al. (Qiu et al., 2022) Epiblast, trophoblast, and extraembryonic endoderm subpopulations were color coded. Lineage differentiation paths between E3.5 and E4.5 were displayed in a tree plot. Right panel displays a force-directed graph showing the trajectory of E4.5 and E5.5 single

cells. (C) Violin plots displaying the transcriptional signature (calculated by UCell) of cell states for E4.5 parietal endoderm module and E4.5 visceral endoderm module identified by hdWGCNA (top 50 genes). n.s. denotes non-significant, **** denotes $p<0.0001$ (Wilcoxon Test). (D and E) Networks connecting the E4.5 parietal endoderm module genes (D) and E4.5 visceral endoderm module genes (E) with enriched KEGG pathways. Individual genes were color coded according to their membership score (kME) relative to the respective gene modules.

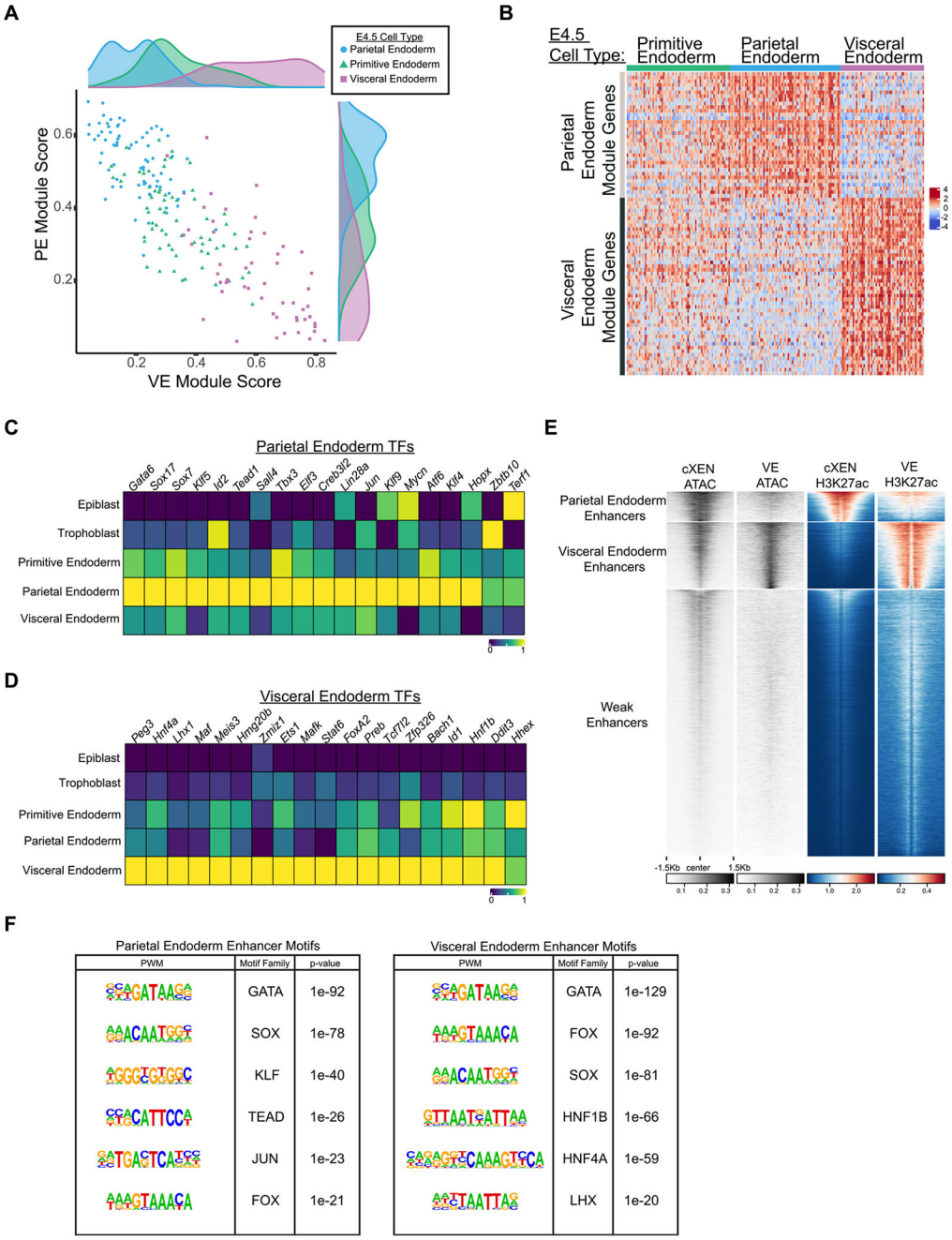


**Fig. 2. Parietal endoderm and visceral endoderm associating transcription factors and enhancers.**

(A) Scatter plot showing the transcriptional signatures of E4.5 parietal endoderm and visceral endoderm gene modules projected onto the single cells of extraembryonic endoderm lineages. Marginal distributions are plotted next to the corresponding axis. (B) Heatmap showing the expression profiles of the top 5% correlated genes (kME > 0) in the parietal endoderm and visceral endoderm module in the single cells of the E4.5 extraembryonic endoderm lineages. (C and D) Average gene expression levels of the top correlated transcription factors in the parietal endoderm (C) or visceral endoderm (D) modules for the

five cell states present in E4.5 embryos. (E) Tornado plots displaying signals of chromatin accessibility and H3K27ac mark in cXEN and E6.5 VE cells centered at H3K27ac enriched regions. (F) Enriched transcription factor binding motifs at parietal endoderm and visceral endoderm enhancers identified in (E).

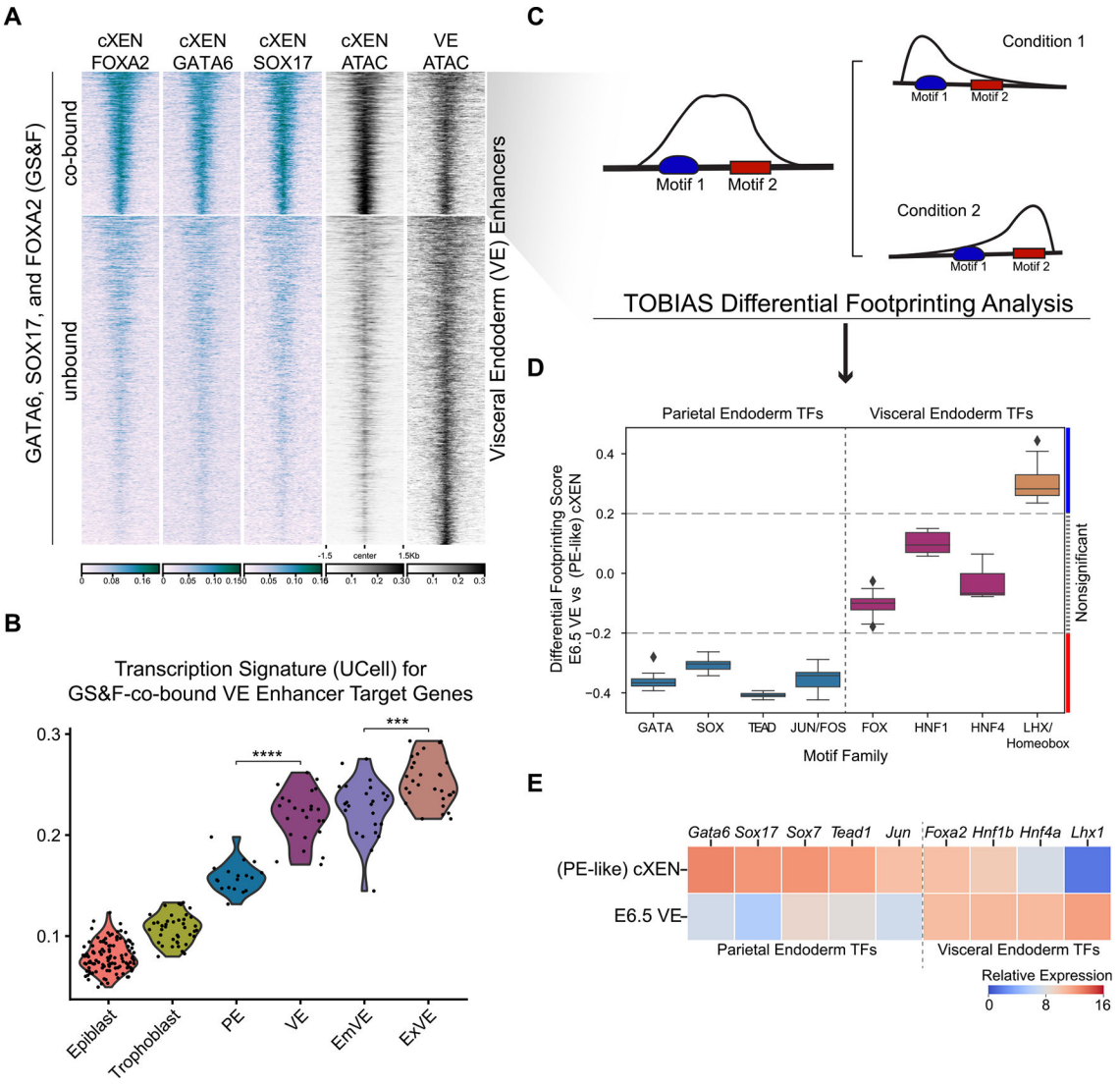


**Fig. 3. Differential transcription factor (TF) footprinting at the GATA6, SOX17, and FOXA2 (GS&F) co-bound VE enhancers.**

(A) Tornado plots displaying signals of chromatin accessibilities and binding of GATA6, SOX17, and FOXA2 around visceral endoderm enhancers identified in Fig. 2E. (B) Violin plot displaying transcriptional signatures of various cell states based on E5.5 scRNA-seq data of differentially expressed genes between E6.5 visceral endoderm and E6.5 epiblast that overlap with the nearby genes of GS&F-co-bound visceral endoderm enhancer identified in (A). *** denotes p<0.001 (Wilcoxon Test) (C) Schematic diagram illustrating differential footprinting analysis using TOBIAS software. (D) Differential footprinting scores for motifs enriched at GS&F-co-bound VE enhancers between cXEN ATAC-seq and E6.5 visceral endoderm ATAC-seq. Similar motifs (PWM matrixes) obtained from the JASPAR and HOCOMOCO databases were clustered into motif families and significance was assigned by TOBIAS, which was visually represented as a boxplot per TF family. Binding scores (y-axis) represent footprint scores calculated by TOBIAS to indicate preferential binding in PE-like (cXEN) cells (i.e. <0 on the y-axis) and preferential binding in E6.5 VE cells

(i.e. >0 on the y-axis). The Log10(p-values) for each PWM matrix can be found in Table S8. Significant differentially bound TFs (indicated by the blue and red color coded bars to the right side of the graph) assigned to cXEN or E6.5 VE cells were defined by default thresholds set by the TOBIAS software. (E) Relative gene expression profiles calculated by DESeq2 median of ratios for TFs belonging to the TF families in (D) from cXEN and E6.5 visceral endoderm RNA-seq datasets.

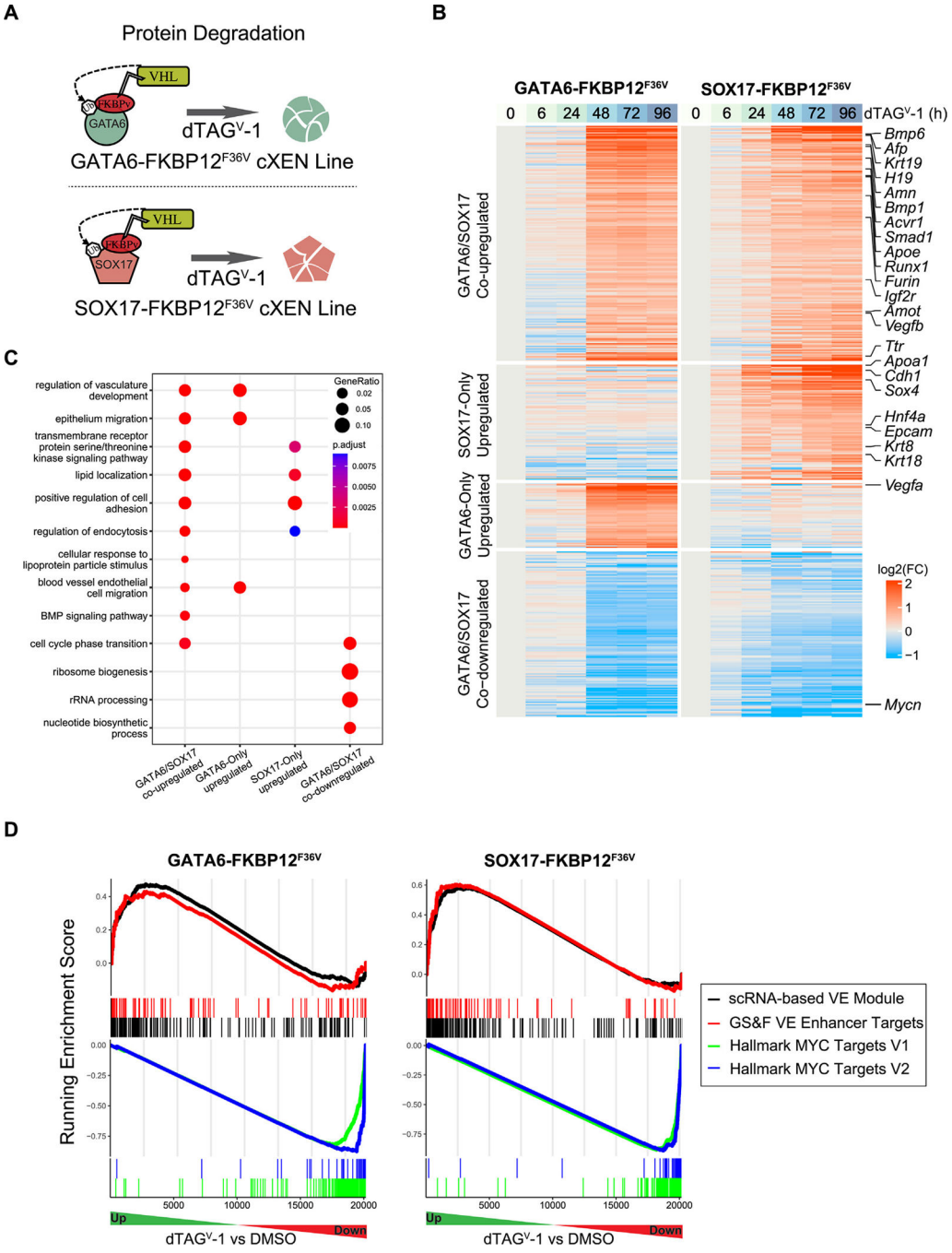


**Fig. 4. GATA6 and SOX17 represses visceral endoderm associated gene expression**

(A) Schematic illustrating the strategy for acute degradation of GATA6 or SOX17 using the dTAG inducible FKBP12$^{F36V}$ degron system (B) Heatmap displaying fold changes (log2 normalized TPM versus 0h) of genes from gene modules identified by WGCNA during the course of GATA6 and SOX17 depletion. Select genes were annotated. GATA6/SOX17 Co-upregulated # of genes: 960, GATA6/SOX17 Co-downregulated # of genes: 681, SOX17 Upregulated # of genes: 469, GATA6 Upregulated # of genes: 265 (C) Aggregated functional profiles of gene ontology biological processes enriched in gene modules in (B)

generated by clusterProfiler. Significance is color coded by p-adjusted values. Gene ratio reflects the number of genes associated with the gene ontology grouping over the total number of input genes. (D) Barcode plot showing gene set enrichment analyses (GSEA) on gene sets (legend on the right) against ranked genes according to expression changes in 72h dTAGV-1 versus DMSO treated *Gata6*-FKBP12$^{F36V}$ or *Sox17*-FKBP12$^{F36V}$ cXEN cells.

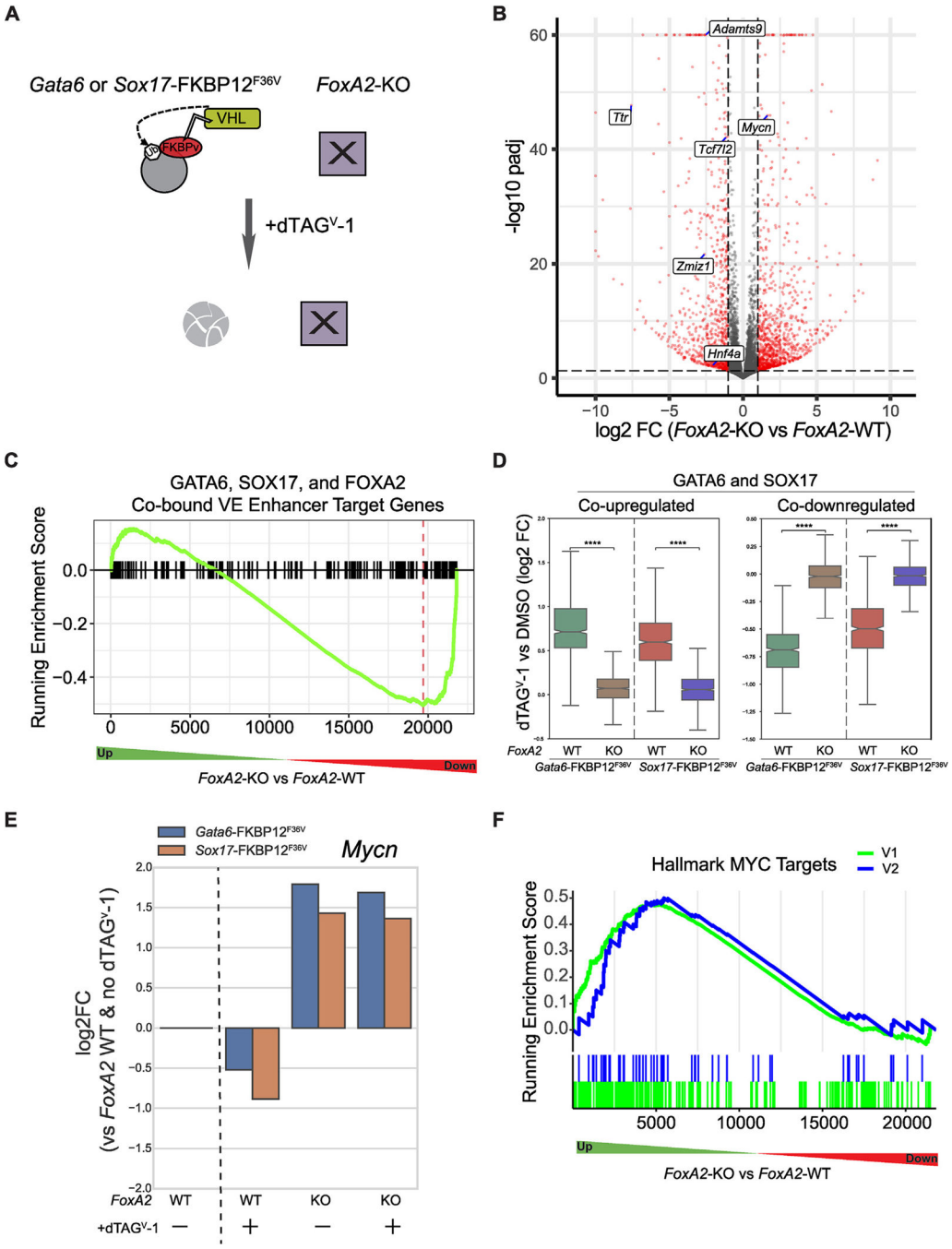


**Fig. 5. FOXA2 counteracts GATA6 and SOX17 in regulating visceral endoderm gene program and *Mycn*.**

(A) Schematic illustrating the strategy for acute degradation of GATA6 or SOX17 (Fig. 4A) in combination with *FoxA2* DNA binding domain deletion. (B) Volcano plot comparing the transcriptomes of *FoxA2*-KO versus *FoxA2*-WT cXEN cells. (C) Barcode plot showing gene set enrichment analyses (GSEA) on GS&F-co-bound VE enhancer targets (Fig. 3B) against ranked genes according to expression changes in *FoxA2*-KO versus *FoxA2*-WT cXEN cells. (D) Boxplot displaying log2 fold changes of GATA6 or SOX17 depletion (dTAG$^{V}$-1 versus DMSO) in *FoxA2*-KO or FoxA2-WT background for GATA6

and SOX17 Co-upregulated and Co-downregulated modules identified in Fig. 4B. **** denotes p<0.0001 (Wilcoxon Test) (E) Barplot displaying fold changes in *Mycn* expression comparing to *Gata6*-FKBP12$^{F36V}$ (blue) or *Sox17*-FKBP12$^{F36V}$ (orange) parental cell lines in the absence of dTAG$^{V}$-1 (first group in the x axis) for *FoxA2*-KO or *FoxA2*-WT backgrounds with and without 48h dTAG$^{V}$-1 treatment. (F) Barcode plot showing gene set enrichment analyses (GSEA) on MYC target genes against ranked genes according to expression changes in *FoxA2*-KO versus *FoxA2*-WT cXEN cells.

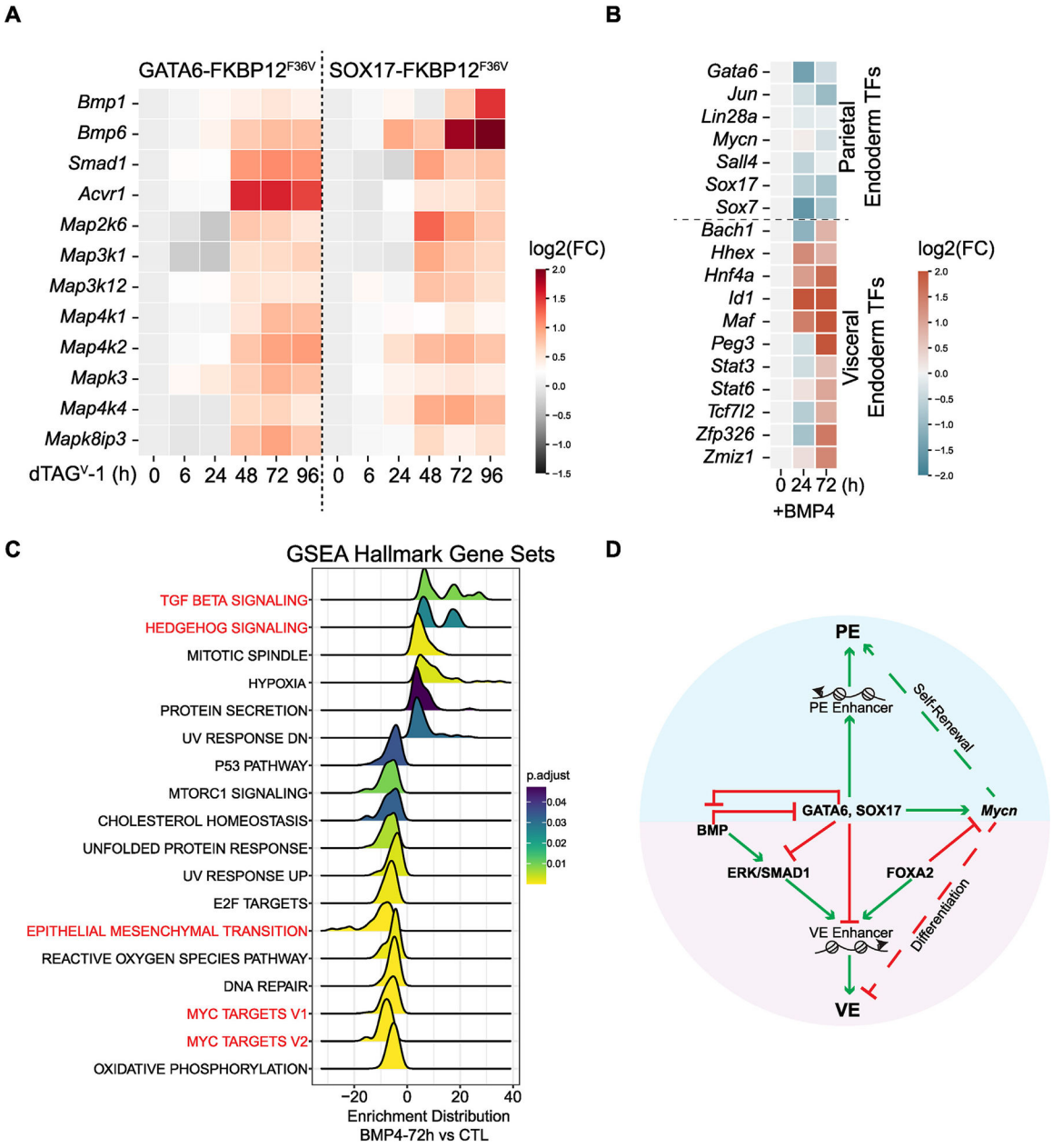


**Fig. 6. Reciprocal inhibition between GATA6 and SOX17 with BMP signaling pathway.**

(A) Heatmap displaying log2 fold changes in normalized TPMs of GATA6-or SOX17-depleted cXEN cells relative to the control for select genes involved in BMP signaling and ERK pathways as annotated in gene ontology databases. (B) Heatmap displaying log2 fold changes in normalized TPMs of BMP4 treated cXEN cells relative to the control for differentially expressed ($q < 0.05$, 72 h vs. 0h DESeq2) TFs in parietal endoderm transcription factor (TF) set (see Fig. 2C) or visceral endoderm TF set (see Fig. 2D). (C) Ridgeplot displaying the distributions of expression changes comparing 72h BMP4 treated cXEN cells versus the control for core genes of the enriched pathways from the MSigDb collection. The significance of pathway enrichments was color coded. (D) Summary of proposed model: GATA6 and SOX17 repress VE gene program and activate

*Mycn* but opposed by FOXA2. Reciprocal repression between GATA6 and SOX17 with BMP signaling modulates the relative expression between PE and VE gene programs.

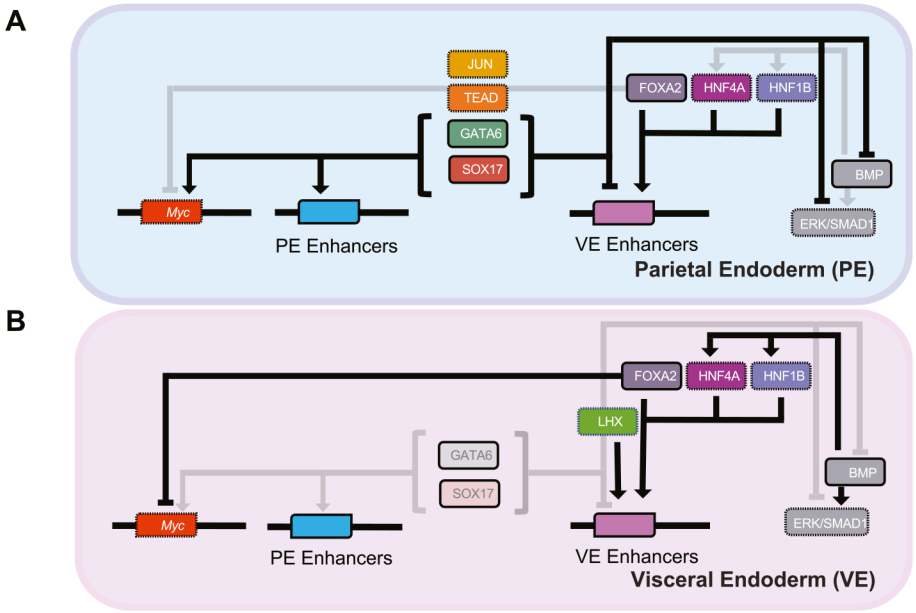


**Fig. 7. Gene regulatory network (GRN) model for PE versus VE cell fate choice.**

(A) and (B) A proposed GRN model for PE versus VE cell fate choice. The model summarizes the hypothesized relationships derived from both scRNA-seq and epigenetic datasets (Figs. 1–3). Boxes with solid black lines represent experimentally tested hypotheses through loss-of-function models (GATA6, SOX17, and FOXA2), and BMP4 treatments using cXEN (PE-like) cells to determine the regulatory nature of these factors underpinning PE versus VE cell fate choice. Boxes with dashed lines represent proposed hypotheses of factors identified through the scRNA-seq and epigenetic analyses, suggesting their potential regulatory nature.